Spatio-temporal analyses of fruit fly populations in selected areas of the Western Cape

by

Gerard Francois Hermanus van Ginkel Bekker

Dissertation presented in fulfilment of the requirements for the degree of

Doctor of Philosophy in Entomology in the

Faculty of AgriSciences at

Stellenbosch University

Supervisor: Dr Pia Addison

Co-supervisor: Mr Matthew Addison

Co-supervisor: Prof Adriaan van Niekerk

April 2019

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated) that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: April 2019

Abstract

The aim of this study was to investigate the spatio-temporal distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) populations in heterogeneous fruit producing environments in Western Cape, South Africa, using geospatial analyses and machine learning (ML) techniques.

A small scale study was conducted at orchard level on the Welgevallen experimental farm in Stellenbosch, South Africa, investigating the spatial patterns and associations of *C. capitata* and *Ceratitis quilicii* females, another important fruit fly pest in the Western Cape. The females of both species had aggregated spatial patterns, but their temporal patterns differed, with *C. capitata* aggregating significantly more towards the end of the season while *C. quilicii* aggregated significantly towards the beginning of the season. *Ceratitis capitata* and *C. quilicii* females were spatially associated, most prominently in home gardens, natural vegetation, citrus and nectarines.

A geographical database was developed, incorporating existing area-wide trap monitoring data for *C. capitata* populations in the Elgin/Grabouw, Villiersdorp, Vyeboom (EGVV) region, Western Cape, an area currently under Sterile Insect Technique (SIT) management, was used to develop a geographical database with the aim to investigate the area-wide spatio-temporal distribution of *C. capitata*. There were no definitive spatial distribution pattern of *C. capitata* across all seasons. However, through visual analyses of spatial maps, a south-east/north-west split was observed where traps in the south-eastern parts of the study area showed higher catches and traps in the north-western parts showed lower catches. The results suggested a relationship between the geographic characteristics of EGVV and the abundance and distribution of *C. capitata* populations.

The relationship between the geographic characteristics of the study area (EGVV) and the spatio-temporal distribution of *C. capitata* were further investigated using ML techniques. Monthly and seasonal long-term *C. capitata* spatio-temporal distributions were quantified into hot-and cold spots (HCSs), using spatial analyses tools. HCSs were then related to a set of geographic variables, using the random forest (RF) ML classification algorithm to determine the main drivers of the HCSs for *C. capitata* in the EGVV region. Spatial analyses

showed that hot spots were concentrated in the hotter and drier areas, while cold spots were concentrated in the colder and wetter areas. The RF results indicated that rainfall was the most important driver of the HCSs in the EGVV region.

To test the robustness of the RF algorithm for the purpose of explaining *C. capitata* HCSs in a heterogeneous fruit producing environment, the sample size and the variability in the geographic variables were increased by combining data from two regions: the EGVV and the Warmbokkeveld (WB), another fruit producing region under SIT. RF model accuracies from the combined dataset were not significantly lower than those of the individual regions. The drivers of *C. capitata* spatial distribution were different between regions, but distance to urban areas in the early fruiting season emerged as a strong driver in all scenarios. The findings showed that RF is a useful tool for investigating the spatio-temporal distribution of area-wide tephritid fruit fly trapping data, and that it can handle complex classification problems. It was evident from this study that the spatio-temporal distribution of *C. capitata* populations are driven by area-specific geographic variables. The area-specific RF models provided invaluable information, which could be used to improve the planning and implementation of area-wide *C. capitata* management programmes in heterogeneous agricultural landscapes.

This study is relevant to the integrated management of fruit flies and potentially other insect pest species, on a local and regional scale. The framework which was developed will allow for the integration of a variety of data and the resultant analyses are relevant at an orchard and regional level. The information will assist efficient decision making by farmers and managers of area-wide integrated pest management programmes

Opsomming

Die doel van hierdie studie was om die tydruimtelike verspreiding van die Mediterreense vrugtevlieg populasies, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae), te ondersoek in heterogene vrugteproduserende omgewings in die Wes-Kaap, Suid-Afrika, deur gebruik te maak van georuimtelike analises en masjienleer (ML) tegnieke.

'n Kleinskaalse studie is eers gedoen op boordvlak op die Welgevallen proefplaas in Stellenbosch, Suid-Afrika. Hierdie studie het beoog om die ruimtelike-patrone en -ooreenstemming van *C. capitata* en *Ceratitis quilicii* te ondersoek. *Ceratitis quilicii* is ook een van die vrugtevlieë in die Wes-Kaap wat ekonomies belangrik is. Resultate het getoon dat die wyfies van beide spesies versamelde ruimtelike-patrone vertoon, maar dat hulle tydspatrone verskil. *Ceratitis capitata* het beduidend aan die einde van die seisoen byeengekom, terwyl *C. quilicii* beduidend aan die begin van die seisoen byeengekom het. Beide spesies se ruimtelike-patrone het ooreengestem, maar die mees prominente ooreenstemming was in huistuine, natuurlike plantegroei, sitrus en nektariens.

Na die bogenoemde studie is 'n geograpfiesedatabasis geskep met die beskikbare area-wye lokvaldata vir *C. capitata* van die Elgin/Grabouw, Villiersdorp, Vyeboom (EGVV) area in die Wes-Kaap. *Ceratitis. capitata* word tans in EGVV bestuur deur gebruik te maak van die Steriele Insek Tegniek (SIT). Die doel van die databasis is om die area-wye tydruimtelike verspreiding van *C. capitata* te ondersoek. Resultate het getoon dat daar geen definitiewe ruimtelike verspreidingspatrone voorgekom het vir die tydperk van al die seisoene wat geanaliseer is nie. Nietemin het die visuele bestudering van kaarte getoon dat daar 'n suid-oos/noord-wes verdeling in lokval-tellings voorgekom het. Lokvalle met hoë tellings was meer verpreid in die suid-oostelike gedeeltes van die studie-area, terwyl die lokvalle met lae tellings meer verpreid was in die noord-westelike gedeeltes van die studie-area. Die resultate stel voor dat daar 'n verhouding bestaan tussen die geografiese karaktereienskappe van EGVV en die oorvloedige voorkoms en verspreiding van *C. capitata*-populasies.

Hierna is die verhouding tussen die geografiese karaktereienskappe van EGVV en die tydruimtelike verspreiding van *C. capitata* verder ondersoek deur gebruik te maak van ML tegnieke. Die maandelikse en seisoenale langtermyn *C. capitata* verspreiding is

gekwantifiseer in warm- en koue kolle (WKKe) deur gebruik te maak van ruimtelike analises. Daarna is ooreenkomstes gevind tussen die WKKe en ’n reeks geografiese veranderlikes deur gebruik te maak van die “random forest” (RF) ML klassifikasie algoritme, met die doel om die hoofdrywers van die WKKe vir *C. capitata* in die EGVV-area te bepaal. Die ruimtelike analises het gevind dat warm kolle in die warmer en droër areas gekonsentreerd was, terwyl die koue kolle meer in die kouer en natter areas gekonsentreerd was. Die RF se resultate het getoon dat reënval die belangrikste drywer, vir die WKKe in die EGVV-streek is.

Om die robuustheid van die RF algoritme te toets met die doel om die *C. capitata* WKKe in heterogene vrugproduserende omgewings te verduidelik, is die steekproef vergroot en die variasie van die geografiese veranderlikes verhoog deur die data van twee streke te kombineer: die EGVV en die Warmbokkeveld (WB), nog ’n vrugproduserende area onder die SIT. Die resultate het getoon dat die RF model van die gekombineerde datastel se akkuraatheid nie beduidend verskil van die akkuraatheid van die individuele areas nie. Die drywers van die ruimtelike verspreiding van *C. capitata* het verskil tussen EGVV en WB, maar afstand vanaf stedelike areas gedurende die vroeë vrugteseisoen het as ’n sterk drywer in beide areas na vore gekom. Die bevindings het gewys dat RF ’n handige hulpbron is om die tydruimtelike verspreding van area-wye vrugtevlieg-lokval data te ondersoek en dat die RF algoritme komplekse klassifikasie probleme kan hanteer. Dit het duidelik getoon dat die tydruimtelike verspreiding van *C. capitata*-populasies gedryf word deur area-spesifieke geografiese faktore. Die area-spesifieke RF modelle het waardevolle inligting verskaf wat aangewend kan word om die beplanning en implimentering van area-wye *C. capitata* bestuursprogramme in heterogene landbou-landskappe te verbeter.

Hierdie studie is relevant tot die geïntegreerde bestuur van vrugtevlieë asook ander potensiële insek-peste op ’n plaaslike- en streeksvlak. Die raamwerk wat ontwikkel is skep die platvorm om ’n verkeidenheid van data-soorte te integreer en te analiseer op boord- en streeksvlak. Boere en bestuurders van area-wye geïntegreerde pes-bestuursprogramme kan die ontginde inligting van die studie gebruik as hulpmiddel in hulle besluitnemingsprosesse.

Dedication

I dedicate this dissertation to

my father and mother, my two brothers and sister, who contributed to my upbringing and to where I am today;

my wife and mother of my children, for her understanding and support throughout this journey;

my twins who were my inspiration and motivation during the dark ending days of this journey, and for making me realise that life is precious; and

my friends and fellow students, who contributed to my personal and academic growth.

Biographical sketch

I grew up on a farm in the Eastern Cape, between Hofmeyr and Steynsburg. This is where I spent my childhood years. I attended primary school from grade one to three in Steynsburg at Paul Kruger College. Thereafter, I went to Cradock Primary School and Cradock High School. I then attended Stellenbosch University to study a BSc in Conservation Ecology. Thereafter, I did a Masters in Philosophy in Environmental Management at the School for Public Leadership at the Stellenbosch University, where the well-known television persona and naturalist, Dave Pepler, supervised me. I spent a year after my masters degree working as a research assistant with the Integrated Pest Management group in the Department of Conservation Ecology and Entomology at Stellenbosch University. This is where I met my good friend and supervisor Dr Pia Addison. I then went on to work for a private firm in Stellenbosch as a Junior Environmental Consultant, doing Environmental Impact Assessments, especially in the agricultural development sector. However, this work did not interest me as much as I would have hoped. Later that year an opportunity arose for me to do my PhD in Entomology under Dr Pia Addison, and I took it with both hands.

I am a keen canoeist; I love bookshops, cooking, the outdoors and my family. I also have a passion for serving and helping others in everyday life.

Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- My supervisor, Dr Pia Addison, and co-supervisor Mr Matthew Addison who believed in me from the word get go. They allowed me to grow independently as a researcher by allowing me the freedom to explore my own ideas.

- My other co-supervisor, Prof Adriaan van Niekerk who is an outstanding scientist that was always available whenever I needed his assistance. He instilled confidence in my academic capabilities and always greeted me with a smile.

- The hard working individuals at the Center for Geographical Analysis (CGA) at the Stellenbosch University, for their speedy technical support, whenever I had a question or problems with my software. I would like to thank Theo Pauw, Divan Vermeulen and Jascha Muller in this regard. Furthermore, to Ilze Boonzaaier for her support in map production

- Prof Daan Nel, Prof Martin Kidd, Dr Ruan Veldtman and Dr Ken Pringle for their assistance and patience with statistical analysis.

- My wife for her love and support throughout this journey, being there when I needed it most, for packing me lunch every day and making sure that there is food when I got home, also for helping me count flies whenever I needed her help.

- Meent Bekker, my brother and my best friend, for being my role model and the one I look up to. Thank you for always listening and giving advice only when asked for, and for giving perspective in times when it was most needed.

- Antoni Schonken, for his willingness to help me with anything, even if he did not know how and his never-ending support.

- Monique James, Saskia Thomas, Bongi Dlamini, Meschack Magagula, Francois du Preez, Pia Addison and Marce Doubell for assisting me with field work.

- Murray Dunn for his weekly assistance with field work and lab work throughout the last year of this study.

- The staff at the Welgevallen Experimental farm: Elmie Lotze, Willem van Kerwel and Basil Titus for their willingness to be of assistance to me and allowing me to use the experimental farm for my experiments.

- Mike Wallace and the Elsenburg GIS department for supplying me with valuable spatial data.

- Nando Baart and the FruitFly Africa staff in Stellenbosch and Grabouw; and Neville Boersma and the XSIT staff in Citrusdal who supplied me with insect monitoring data.

- People who have supported me with ideas, and with whom I could always have a conversation about my work, whenever I needed to: thank you to Cory Bazelet, Aruna Manrakkan, Marc De Meyer, Minette Karsten, Marce Doubell, Craig Galloway, Ruan Veldtman, Maulid Mwatawala and Caro Kapp.

- Oom Johann Adler for being there to listen and ask questions, whenever life through its challenges towards me.

- Helene Van Niekerk for proof reading parts of this dissertation.

My sincere gratitude goes out to the funders of this research: HortGro POME, HortGro STONE, Stellenbosch University, The International Atomic Energy Agency (IAEA) and the National Research Foundation (NRF).

Preface

This dissertation is presented as a compilation of six chapters.

Chapter 1		General introduction
Chapter 2		Spatial patterns and associations of *Ceratitis capitata* and *Ceratitis quilicii* (Diptera: Tephritidae) females in a small-scale heterogeneous landscape in the Western Cape, South Africa
Chapter 3		The area-wide spatio-temporal distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) in a heterogenous fruit production region of the Western Cape, South Africa
		This chapter has been accepted as a publication in *African Entomology*
Chapter 4		Using machine learning to identify the geographical drivers of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) trap catch in an agricultural landscape
		This chapter has been accepted as a publication in *Computers and Electronics in Agriculture*
Chapter 5		Investigating the robustness of the random forest machine learning algorithm to classify *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) population hot- and cold spots in two fruit producing regions of the Western Cape, South Africa
Chapter 6		General discussion and conclusions

Chapters 2 to 4 are written as separate research papers and, for that reason, a certain amount of repetition was unavoidable.

Chapter 5 is written as a separate dissertation chapter expanding on Chapter 4.

Table of Contents

Chapter 1

General introduction

Introduction

About 250 tephritid fruit fly species (Diptera: Tephritidae) of the more than 4000 species of Tephritidae fruit flies found worldwide, are known to attack numerous fruits types that are grown commercially or harvested in the wild (White & Elson-Harris 1992). These fruit flies cause excessive economic losses to horticultural industries in many regions of the world, which place tremendous pressure on food security, especially in countries where fruit and vegetables are the main source of nutrition and income (Allwood & Drew 1996; Ekesi 2010).

Tephritids are widely distributed and occur in almost all temperate, subtropical and tropical regions of the world (Christenson & Foote 1960). Almost all species are phytophagous, with larval development that can take place within fruits, flowers, or the stems of host plants (De Meyer *et al.* 2008). Depending on climate conditions and abundance of host fruits, fruit flies may complete several generations in a year. The general life cycle of tephritid fruit flies are displayed in Figure 1.

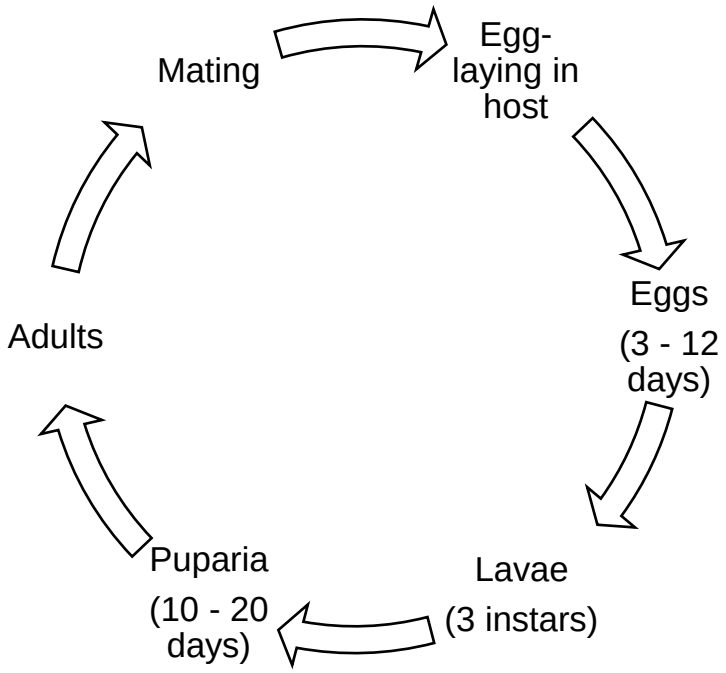

Figure 1. A generalised life cycle of tephritid fruit flies (adapted from Ekesi & Billah 2007).

In Africa, native and invasive fruit flies cause direct fruit damage as well as indirect damage through quarantine restrictions enforced by importing countries (Ekesi 2010). Such quarantine restrictions can either deny the producing country from exporting to certain markets or force local producers to comply with expensive management and control practices (White & Elson-Harris 1992). This has implications for the socio-economic landscape of the continent, because it can lead to food insecurity, job losses, loss of income and employment in certain cases (Ekesi 2010). In some African countries such as Benin, Tanzania and Ghana direct crop damages of up to 85% have been reported due to *Bactrocera dorsalis* Hendel (Diptera: Tephritidae) attacking fruits (Ekesi *et al.* 2006; Mwatawala *et al.* 2006; Vayssières *et al.* 2008). In Mozambique, the temporary quarantine restrictions that were placed on fruit and vegetable exports to South Africa in 2008, due to the presence of *B. dorsalis*, resulted in monitory losses of up to two million US dollars (Cugala 2011).

In South Africa, the export of fresh fruits is a major foreign currency earner. In 2016, the total turnover of the deciduous fruit industry was approximately 13 billion Rand and 44% of the total South African deciduous fruit production were exported (Hortgro 2016). The deciduous fruit industry also provided 1,34 permanent jobs per hectare in 2016. In terms of deciduous fruit production, the Western Cape province outweighs the other provinces, in South Africa (Hortgro 2016). In excess of 50 000 ha of deciduous fruit orchards and vineyards are cultivated in the Western Cape out of a total of 80 000 ha of deciduous fruit orcahrds and vineyards in South Africa (Hortgro 2016). These statistics indicate that not only the South African deciduous fruit industry as a whole, but more specifically the Western Cape deciduous fruit industry, is of significant socio-economic importance to the country.

Economic cost of fruit flies to the South Africa deciduous fruit industry

There are five fruit fly species in South Africa that are of significant economic importance, namely *Ceratitis capitata* (Wiedemann), the Mediterranean fruit fly (Medfly), *Ceratitis rosa* Karsch, the Natal fruit fly, *Ceratitis quilicii* De Meyer, Mwatawala and Virgilio, the Cape fruit fly, *B. dorsalis* (Hendel), the Oriental fruit fly, and *Ceratitis cosyra* Walker, the mango fly (Annecke & Moran 1982; Barnes *et al.* 2002; Blomefield *et al.* 2015; Manrakhan *et al.* 2015; Karsten *et al.* 2018). *Ceratitis capitata* is the most widely spread and regarded as the most important to the deciduous fruit industry (DeVilliers *et al.* 2013). In 1997, *C. capitata* alone was estimated to cost the Western Cape deciduous fruit industry roughly more than R40 million per annum in crop losses and control costs (Mumford & Tween 1997). This excluded the indirect costs that can be associated with quarantine restrictions, preventing producers from exporting their fruit to lucrative markets around the world. No recent monetary estimation has been calculated, but fruit flies still remain one of the major fruit pests in the country today (Badii *et al.* 2015; Barnes *et al.* 2015).

Control of fruit flies

Chemical control using insecticidal cover sprays has previously been the principle method of fruit fly control (Rössler 1989). Inorganic insecticide cover sprays in the form of lead arsenate and sodium fluorosilicate were commonly used in the early 1900's (Joubert 1932). Organophosphates were introduced later, of which Malathion has been used successfully to control *C. capitata* populations throughout the world (Rössler 1989).

Currently deciduous and citrus fruit producers in South Africa, use a number of different control techniques against fruit flies. The first includes full cover sprays using insecticides (organophosphates), of which the residues can have negative effects on human and environmental health. The second method is the bait application technique (BAT), which consists of a mixture of a bait/attractant (normally a hydrolysed protein) and a pesticide (for example organophosphates, spinosad, or synthetic pyrethroids) which attracts and then kills the fruit fly (Roessler 1989). BAT has an advantage over cover sprays in that it is used to apply spot treatments away from the fruit, thus minimising the negative effect of chemical residues on human health and the environment (Stonehouse *et al.* 2002; Prokopy *et al.* 2003; Vayssières *et al.* 2009). The bait mixture used in BAT is sometimes placed in a container (for example a trap) and hung from a tree to limit the insecticidal exposure, but also to retain the flies within the container (Ekesi & Billah 2007). Another, less costly, trapping alternative is bait blocks, which consists of a piece of absorbable material, such as compressed fibre board, which is soaked in bait mixture and nailed to a tree or row post to attract and kill the flies (Ekesi & Billah 2007).

All methods still use some form of insecticides, which can lead to resistance to insecticides and that are harmful to the environment and human health, therefore there has been a major drive, due to consumer demand, towards more environmentally friendly control methods, of

which the Sterile Insect Technique (SIT) is one. The concept of SIT entails the mass rearing and sterilization of males, releasing the sterilised males into the field to mate with wild females, preventing them from producing fertile eggs (Knipling 1960; Gilmore 1989), thus acting as a type of birth control. SIT is generally most effective when the wild population is low, thus it is most effective when being used in combination with other control methods, such as BAT, biological control and sanitation, which is used to reduce the target populations to levels at which SIT is effective (Gilmore 1989).

The integration of control methods such as SIT and BAT could be seen as an integrated pest management (IPM) strategy. In a review article, Kogan (1998), defined a formal definition of IPM, which reads:

> "IPM is a decision support system for the selection and use of pest control tactics, singly or harmoniously coordinated into a management strategy, based on cost/benefit analyses that take into account the interest of and impacts on producers, society, and the environment."

Because of the mobility of fruit flies and their ability to distribute over large distances (Fletcher 1989), area-wide integrated pest management (AW-IPM) strategies have been proposed to effectively manage fruit flies in large heterogeneous agricultural production areas (Klassen 2005; Hendrichs *et al.* 2015a). Klassen (2005) defined AW-IPM as:

> "IPM against an entire pest population within a delimited geographic area, with a minimum size large enough or protected by a buffer zone so that natural dispersal of the population occurs only within this area."

However, for the purpose of this dissertation the definition by Lindquist (2000) for AW-IPM better describes the situation in the Western Cape, South Africa, which states that:

> "An area-wide insect control programme is a long-term planned campaign against a pest population in a relatively large predefined area with the objective of reducing the insect pest population to a non-economic status."

Fruit fly control in the Western Cape, South Africa

Currently a combination of BAT and SIT is used in the major deciduous fruit production regions of the Western Cape to control fruit flies on an area-wide basis (Barnes *et al.* 2015). In South Africa, SIT for area-wide application is only available in the Western Cape and only for *C. capitata.*

The area-wide control of fruit flies in the Western Cape is conducted as part of an area-wide integrated pest management programme (AW-IPM), which is operated and managed by a private non-profit company, namely FruitFly Africa (Pty) Ltd. (FFA). Their responsibilities include a wide variety of tasks pertaining to fruit fly management. Their main tasks, however, are the mass-rearing, sterilisation and routine releases of sterilised males into target areas, the coordination and supervision of area-wide BAT in target areas and the monitoring of wild and sterile fruit fly populations (FruitFly Africa 2017). The detailed history and the development of this programme is outlined in Barnes *et al.* (2015).

For any control measure or management programme to be applied effectively towards any pest, two questions need to be answered. The first being, where is the pest located within the landscape or on the plant? The second being when does this pest occur in time? These two questions put space and time into the equation of deciding when, where and what control measures or management actions need to be taken. Therefore, knowing the spatio-temporal distribution of a pest is of utmost importance for effective pest control or management (Nestel *et al.* 2002; Hendrichs *et al.* 2007b; Kounatidis *et al.* 2008; Sciarretta & Trematerra 2011; Pimentel *et al.* 2014). This is also true for applying control measures and management

actions towards fruit flies. Knowing the spatio-temporal distribution of a pest is important, however, probably more important is to know why a pest displays a certain spatio-temporal distribution or pattern. To answer this question it is important to know what abiotic or biotic factors might influence insect distributions.

In any integrated fruit fly management programme, whether conducted on an area-wide scale or at an orchard level, host management and sanitation are important management actions to implement, to prevent population build-up (Barnes 2009; Vargas *et al.* 2010). Host management consists of removing fruits from the tree (fruit stripping) or applying an insecticide, to manage fruit fly populations. After fruit stripping, the next step is sanitation, which involves destroying all possible fruit fly eggs and larvae in the fruit through pulping, burying the fruit underground (± 1 m) or placing the fruit in a plastic bag and leaving it in the sun (Barnes 2009). Sanitation can also be conducted by using augmentoria (Klungness *et al.* 2005; Jang *et al.* 2007), allowing fruit fly parasitoids to escape the tent-like structure, but killing flies within, and thereby contributing to fruit fly management. According to Barnes (2009), unmanaged home garden hosts, neglected orchards and vineyards and alternate hosts in urban areas and natural vegetation all contribute substantially to fruit fly populations within an area. Although host management and sanitation are recommended to farmers in South Africa and the Western Cape, it is difficult to quantify, and this quantification towards establishing the value of the practice has not been done in South Africa. However, studies in Hawaii showed a significant impact of augmentoria on the emergence of fruit flies (Klungness *et al.* 2005; Jang *et al.* 2007).

Factors effecting the spatio-temporal distribution of fruit flies

Biotic, abiotic, and anthropogenic factors have been shown to influence the spatio-temporal dynamics and distribution of organisms within a landscape. Results from Nestel *et al.* (2002)

suggest that the use of pesticides, that can be seen as an anthropogenic factor, plays a role in shaping the spatio-temporal distribution of fruit flies in some areas. For fruit flies to survive they need essential resources such as food, mates, oviposition sites and refuge (Prokopy *et al.* 1994). Thus, it is safe to assume that fruit flies will respond to the spatio-temporal and seasonal distribution of these resources.

In numerous studies investigating fruit fly distributions, it has been found that fruit fly spatial patterns are influenced by the availability and distribution of host fruits (Rivnay 1954; Vargas *et al.* 1983b; Katsoyannos *et al.* 1998; Papadopoulos *et al.* 2001b, 2003; Nestel *et al.* 2002; Alemany *et al.* 2006; Sciarretta & Trematerra 2011). Furthermore, Pimentel *et al.* (2014) suggest that the spatial patterns of *C. capitata* might be influenced by the microclimate that exists within the different topographical regions of an area. Results from DeVilliers *et al.* (2013) suggest that temperature, humidity and rainfall are the most important factors that influence the abundance and distribution of *C. capitata* and *C. rosa* in South Africa. This corresponds with conclusions drawn from other studies that suggest that climate plays a major role in *C. capitata* distribution and abundance (Myburgh 1962; Duyck *et al.* 2006b; Nyamukondiwa & Terblanche 2009, 2010). In line with these conclusions, Peñarrubia-María *et al.* (2012) found that rainfall influenced the abundance of *C. capitata* in North Eastern Spain, thus also having an effect on the distribution of flies.

From the referenced studies above, it is clear that climatic factors, geographical factors, host fruit availability and distribution, but also anthropogenic factors such as pesticide applications, influences and shapes the spatio-temporal distribution of fruit flies within a landscape.

The use of traps in area-wide fruit fly management programmes

Monitoring pest populations using traps, is a relative sampling method, only giving an estimate of the population abundance as opposed to a direct sampling method that will give an absolute measure (Midgarden *et al.* 2014). With a variety of traps and attractants, each developed with its own goal in mind, makes most trapping systems biased (Vale & Phelps 1978; Vreysen & Saleh 2001; Ekesi *et al.* 2005). Therefore, when interpreting results, trap biases should be considered. Some of the most significant factors that can influence trap catch, listed by Vreysen (2005), include trap efficiency and insect behaviour.

Trapping efficiency or effectiveness has been shown to be influenced by the shape, size and colour of the trap (Eliopoulos 2007; Navarro-Llopis *et al.* 2008; Robacker 2015). Robacker *et al.* (1990) studied the effect of trap placement on captures and found it to be a significant factor to consider. A range of climatic conditions including temperature, rainfall, relative humidity and light intensity can influence trapping efficiency (Haniotakis 1974; Kapatos & Fletcher 1983; Drake 1994; Kitron 1998). Furthermore, depending on the behaviour and the physiological state of an organism, the likelihood of an organism responding to a trap can be variably influenced (Neuenschwander & Michelakis 1979; Vreysen & Saleh 2001; Coracini *et al.* 2004).

To study the spatio-temporal distribution of fruit flies, a variety of different trap types that vary in shape, size, colour and design, are used. The decision as to which trap to use depends on the species and research objective. Data from fruit fly traps are generally used to enhance our understanding of the behaviour and ecology of these organisms (Midgarden *et al.* 2014). Therefore, traps are used for making decision regarding planning and implementation of management actions in AW-IPM programmes (Cox & Vreysen 2005; Klassen 2005).

Knowing where and when the pest occurs within an area is vital information for AW-IPM programme managers, in order to apply the most efficient control methods at the right location at the right time, which can ultimately contribute to the reduction in programme costs and effective management of the target pest (Cox & Vreysen 2005).

Traps are mostly used in fruit fly AW-IPM programmes for two reasons: 1) to determine wild population size estimates and 2) to determine the spatio-temporal distribution of wild fly populations within the landscape (Midgarden *et al.* 2014). When SIT is part of an AW-IPM programme, traps are also used to determine the spread (dispersal and mobility) and distribution of sterile flies, and to monitor sterile to wild male ratios (Vreysen 2005). This information is useful to programme managers to evaluate the effectiveness of their control techniques. A sterile male to wild fly ratio of at least 80:1 (sterile males:wild males) is recommended for SIT against *C. capitata* to be fully effective (Barnes 2009). Enkerlin *et al.* (2016) found that when sterile male to wild male ratios were maintained at 100:1, more effective control of *C. capitata* was achieved, compared to lower ratios. Enkerlin *et al.* (2016) used 10 years of trapping data obtained from the Guatemala-Mexico-United States Mediterranean fruit fly Containment and Eradication Programme (Moscamed Programme).

Most AW-IPM programmes including those that incorporate SIT, rely on resource intensive trapping grids (Midgarden *et al.* 2014; Enkerlin *et al.* 2015), that have to be maintained and monitored on a regular basis. These traps generally consist of a container baited with an attractant (parapheromones or food bait) and killing agent to retain flies. Extensive trapping grids used for fruit fly population monitoring, as part of AW-IPM programmes, generate vast amounts of trapping data. Trapping data obtained from these trapping grids are information-rich and could be used to conduct in-depth analyses that can provide valuable information

to help guide decision makers to improve their programmes (Midgarden *et al.* 2014). Generally, AW-IPM trapping surveys not only gather information regarding the location of the pest, but also collect attribute location data that relates to host plants (fruit availability and suitability), trap type, and time which can be valuable for research.

Very little research has been done on the spatio-temporal distribution of fruit flies using trapping data from AW-IPM programmes (see Castrignanò *et al.*, 2012; Guidotti *et al.*, 2005; Kounatidis *et al.*, 2008; Midgarden and Lira, 2006). This is mainly due to the fact that these datasets are complex regarding the spatio-temporal scales at which sampling was conducted, but also because the sampling regimes are guided by programme objectives and not necessarily research questions (Midgarden *et al.* 2014).

The use of geographic information systems in AW-IPM programmes

The development and evolution of geographic information systems (GIS), global positioning systems (GPS) and remote sensing (RS) has facilitated and simplified the collection, integration and analysis of spatial data in recent years, across multiple fields of study. A GIS is a computer-based system that allows its users to capture, integrate, store, retrieve and most importantly, display spatial data onto easy to understand and interpretive maps (Liebhold *et al.* 1993; Cox and Vreysen 2005). Remote sensing is the process of gathering information about an object from a distance. This process is generally carried out by satellites or aircraft collecting data from the earth's surface using a range of electromagnetic sensors attached to these vehicles (Dalsted 2011).

GIS is especially useful for entomological research in that it enhances the ability to study and understand the large-scale spatial structure and dynamics of insect populations in heterogeneous landscapes (Dminic *et al.* 2010). It makes provision for spatial analysis,

which can be described as manipulating and transforming spatial data in such a manner that additional information can be extracted from the data (Bailey 1994; Cox & Vreysen 2005). GIS software also allows for the characterization and modelling of spatial patterns pertaining to spatial insect pest data (trapping data), making use of geostatistics (Liebhold *et al.* 1993; Midgarden *et al.* 1993; Ribes-Dasi *et al.* 2001; Lyons *et al.* 2002; Sciarretta and Trematerra 2006). Geostatistics is a set of statistical tools that can be used to explore and describe spatial dependence among samples and predict values at unsampled locations (Curran & Atkinson 1998; Sciarretta & Trematerra 2006; Dale & Fortin 2014). Geostatistics is an extension of spatial analysis used to investigate spatial patterns and relationships within spatial data.

GIS and geostatistics have been used in numerous studies to investigate the spatial distribution of insect pests (Ribes-Dasi *et al.* 2001; Lyons *et al.* 2002; Trematerra *et al.* 2004; Sciarretta & Trematerra 2006). Ribes-Dasi *et al.* (2001) used GIS and geostatistics to study the spatial distribution of *Cydia pomonella* Linnaeus (Lepidoptera: Tortricidae) and *Pandemis heparana* Denis & Schiffermüller (Lepidoptera: Tortricidae), using pheromone trap catches. Ribes-Dasi *et al.* (2001) identified areas that have a high risk of pest attack and determined the optimal number of traps, to use, and their locations. In Italy, Sciarretta and Trematerra (2006) characterised the spatial distribution of two agricultural pests within a mixed orchard landscape, using pheromone trap catches and analysed them using GIS and geostatistics. Papadopoulos *et al.* (2003) used GIS and geostatistics, more specifically the Moran *I* spatial statistic (Moran 1948), to analyse the spatial autocorrelation (correlation of a variable with itself through space) between *C. capitata* populations at sampling locations and dates to study the spatial dispersion of these flies in the early season. They used pheromone (male biased) and synthetic food based lure (female biased) traps for their analysis within a mixed fruit orchard. Kounatidis *et al.* (2008) made use of the Gertis-Ord Gi*

local spatial statistic methodology (Getis & Ord 1996), to investigate the spatial patterns of tephritid fruit flies in northern Greece. Kounatidis *et al.* (2008) were successful in classifying trap locations into classes of hot spots, were traps with high catches clustered in space, and cold spots, areas where trap with low trap catches clustered in space.

This information is useful as it can help to establish management priority areas. Furthermore, the statistical classification of such spatial patterns presents the opportunity to explore the underlying spatial processes and or factors that may have caused or influenced the spatial patterns displayed. A detailed account of how this methodology works can be found in Kounatidis *et al.* (2008) and Getis and Ord (1992).

There are numerous other studies employing GIS, spatial analysis and geostatistics to investigate the spatio-temporal distribution of *C. capitata* and investigate the reasons for their spatial-temporal distributions displayed (Alemany *et al.* 2006; Epsky *et al.* 2010; Sciarretta & Trematerra 2011). Although GIS, spatial analysis and geostatistics are commonly used in insect pest spatial distribution studies, most of these studies were conducted on orchard or farm level but not at regional scales.

The use of GIS and its integrated tools (spatial analysis and geostatistics) have been used in some of the most successful AW-IPM programmes against major insect pests. These include *C. capitata* and New World screwworm *Cochliomyia hominivorax* Coquerel (Diptera: Calliphoridae) in the Americas (Wyss 2000; Cox & Vreysen 2005) and tsetse fly *Glossina austeni* Newstead (Diptera: Glossinidae) in Zanzibar, Tanzania (Vreysen 2000). The main uses of these technologies in *C. capitata* AW-IPM programmes include: overlaying topographical layers over satellite imagery to identify appropriate trapping sites; and associating wild trap catches with host distribution and evaluating sterile *C. capitata* male

releases by displaying wild to sterile *C. capitata* ratios on a map. However, the AW-IPM programme against *C. capitata* in Guatemala (Moscamed) used GIS and the large available trap monitoring datasets to test sterile fly performance pertaining to a range of abiotic factors such as elevation and habitat. They have also used it to identify and investigate areas where wild *C. capitata* populations persist even when control methods are applied (Cox & Vreysen 2005). Midgarden *et al.* (2014) used GIS and data from the Moscamed programme to study the ecological relationships between *C. capitata* and coffee in Guatemala and Mexico. Kounatidis *et al.* (2008), employed GIS and spatial statistics on data generated by an area-wide trapping network for *Bactrocera oleae* (Rossi) (Olive fruit fly) in Greece, investigating the effect of elevation on spatio-temporal patterns, while Guidotti *et al.* (2005) used it to predict the start of *B. oleae* fruit infestations in orchards.

A relatively new, but fast growing approach towards understanding and exploring complex relationships in large ecological data sets is machine learning (ML) (Olden *et al.* 2008). ML is a field within computer science which involves creating computer algorithms which can be trained (with data) to perform certain tasks and solve complex problems by learning from experience (Mitchell 2006). Machine learning methods have been widely used in the fields of engineering (Aytug *et al.* 1994), medicine (Jothi *et al.* 2015), advertising (Broder *et al.* 2007),economics (Carbonneau *et al.* 2008), artificial intelligence (Hamet & Tremblay 2017) and forestry (Demertzis *et al.* 2015). For an overview on ML please refer to Carbonell *et al.* (1983). Jordan & Mitchell (2015), discusses the latest trends, perspectives and future applications of ML. Although ML has been widely used in numerous fields, its application in ecological studies has been relatively limited. According to Olden *et al.* (2008) this can be ascribed to ML techniques being still largely unfamiliar to ecologists compared to the traditional statistical modelling approaches that they are used to. One of the subjects where it shows great potential is the field of species distribution modelling (Elith & Leathwick 2009).

Some reasons that makes the use of ML methods attractive to use in ecological studies is that it possess the ability to model complex, non-linear relationships in ecological datasets without the need to meet limiting assumptions associated with traditional parametric modelling approaches (Guisan & Zimmermann 2000; Olden & Jackson 2002).

Research Gaps

In South Africa, the use of GIS and associated integrated tools have seen limited use in investigating the spatio-temporal distribution of fruit flies on orchard, farm or regional scale and therefore gaining insights into which factors drives these spatio-temporal distributions, are important, for better management and control of fruit flies. Thorough analysis of trapping data, in space and time, which are generated by resource intensive trapping surveys, which are part of AW-IPM programmes, have not been explored. This reveals the underutilization of information-rich data on which vast amounts of resources are spent every year. With the capabilities that are presented by GIS and ML technologies to easily organise, explore and analyse large complex ecological datasets, the door opens to utilise trapping datasets, generated from AW-IPM programmes. These tools can contribute to sketching a more realistic picture of the spatial relationships and patterns that exist between the target pest and its environment.

The following research questions are of relevance:

- What are the spatial distribution patterns of fruit flies in a heterogeneous orchard environment that are under intensive fruit fly management?
- What are the spatial distribution patterns of fruit flies in heterogeneous fruit growing regions that are under AW-control?
- Can geographic characteristics of regions explain the occurrence of spatial clusters (hot spots) or areas of dispersal (cold spots)?

- Can biotic and abiotic factors be quantitatively linked to these hot- and cold spots?

Aim and objectives of this study

The aim of this study was to investigate the spatio-temporal distribution of *C. capitata* populations on a small scale in a heterogeneous orchard environment and on a regional scale in two large heterogeneous fruit production areas in the Western Cape, using geospatial analyses and ML. The objectives of the study were to:

- Investigate the spatial patterns and associations of *C. capitata* and *C. quilicii* (Diptera: Tephritidae) females in a mixed fruit farm in the Western Cape, South Africa
- Develop a geographical database incorporating existing area-wide trap monitoring data for *C. capitata* populations in the Western Cape, South Africa,
- Investigate the area-wide spatio-temporal distribution of *C. capitata* populations within a heterogeneous fruit production area;
- Model the area-wide spatial distribution of *C. capitata* population hot spots making use of ML classifiers;
- Evaluate the robustness of ML fruit fly models.

Note that this study was exploratory in nature and no hypothesis testing was conducted.

References

ALEMANY, A., MIRANDA, M.A., ALONSO, R. & MARTÍN ESCORZA, C. 2006. Changes in the spatial and temporal population density of the Mediterranean fruit fly (Diptera: Tephritidae) in a citrus orchard. *Spanish Journal of Agricultural Research* **4**(2): 161–166.

ALLWOOD, A.J. & DREW, R.A.I. 1996. Management of Fruit Flies in the Pacific. In: Allwood, A.J. & Drew, R.A.I. (Eds.) *Management of Fruit Flies in the Pacific*. 267. Nadi, Fiji.

ANNECKE, D.P. & MORAN, V.C. 1982. *Insects and Mites of Cultivated Plants in South Africa.* Butterworth, Durban.

AYTUG, H., BHATTACHARYYA, S., KOEHLER, G.J. & SNOWDON, J.L. 1994. A Review of Machine Learning in Scheduling. *IEEE Transactions on Engineering Management* **41**(2): 165–171.

BADII, K.B., BILLAH, M.K., AFREH-NUAMAH, K., OBENG-OFORI, D. & NYARKO, G. 2015. Review of the pest status, economic impact and management of fruit-infesting flies (Diptera: Tephritidae) in Africa. *African Journal of Agricultural Research* **10**(12): 1488–1498.

BAILEY, T.C. 1994. A review of statistical spatial analysis in geographical information systems. In: Fotheringham, S. & Rogerson, P. (Eds.) *Spatial Analysis and GIS*. 13–44. Taylor & Francis, London.

BARNES, B.N. 2009. The Sterile Insect Technique (S.I.T.) for management of Mediterranean Fruit Flies in fruit production - A technical manual. Stellenbosch, South Africa.

BARNES, B.N., EYLES, D.K. & FRANZ, G. 2002. South Africa's fruit fly SIT programme - the Hex River Valley pilot project and beyond. *6th International Fruit Fly Symposium*. 131–141. Stellenbosch.

BARNES, B.N., HOFMEYR, J.H., GROENEWALD, S., CONLONG, D.E. & WOHLFARTER, M. 2015. The sterile insect technique in agricultural crops in South Africa : a metamorphosis .... but will it fly ? *African Entomology* **23**(1): 1–18.

BLOMEFIELD, T.L., BARNES, B.N. & GILIOMEE, J.H. 2015. Peach and nectarine. In: Prinsloo, G.L. & Uys, V.M. (Eds.) *Insects of Cultivated Plants and Natural Pastures in Southern Africa*. 320–339. Entomological Society of Southern Africa, Hatfield.

BRODER, A., FONTOURA, M., JOSIFOVSKI, V. & RIEDEL, L. 2007. A semantic approach to contextual advertising. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. 559–566. ACM.

CARBONELL, J.G., MICHALSKI, R.S. & MITCHELL, T.M. 1983. An overview of machine learning. In: Michalski, R.S., Carbonell, J.G. & Mitchell, T.M. (Eds.) *Machine learning: An Artificial Intelligence Approach*. 3–23. Springer, Palo Alto, California.

CARBONNEAU, R., LAFRAMBOISE, K. & VAHIDOV, R. 2008. Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research* **184**(3): 1140–1154.

CASTRIGNANÒ, A., BOCCACCIO, L., COHEN, Y., NESTEL, D., KOUNATIDIS, I., PAPADOPOULOS, N.T., DE BENEDETTO, D. & MAVRAGANI-TSIPIDOU, P. 2012. Spatio-temporal population dynamics and area-wide delineation of *Bactrocera oleae* monitoring zones using multi-variate geostatistics. *Precision Agriculture* **13**(4): 421–441.

CHRISTENSON, L.D. & FOOTE, R.H. 1960. Biology of Fruit Flies. *Annual Review of Entomology* **5**(1): 171–192.

CORACINI, M., BENGTSSON, M., LIBLIKAS, L. & WITZGALL, P. 2004. Attraction of codling moth males to apple volatiles. *Entomologia Experimentalis et Applicata* **110**(1): 1–10.

COX, J.S.H. & VREYSEN, M.J.B. 2005. Use of Geographic Information Systems and Spatial Analysis in Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 453–477. Springer, Dordrecht.

CUGALA, D.R. 2011. Surveillance, Management and Control strategies of the invasive fruit fly *Bactrocera invadens* in Mozambique. *Workshop presentation about Fruit flies in Mozambique. Pemba 23-24th March, 2011*.

CURRAN, P.J. & ATKINSON, P.M. 1998. Geostatistics and remote sensing. *Progress in Physical Geography* **22**(1): 61–78.

DALE, M.R.T. & FORTIN, M.-J. 2014. *Spatial Analysis: A Guide for Ecologist* (2nd edition). Cambridge University Press, Cambridge, United Kingdom.

DALSTED, K. 2011. Introduction: Remote Sensing and GIS Techniques for the Detection, Surveillance, and Management of Invasive Species. Volume 3: Invasive Species. In: Clay, S.A. (Ed.) *GIS Applications in Agriculture*. 1–8. CRC Press, Taylor & Francis Group, Boca Raton.

DEMERTZIS, K., ILIADIS, L., AVRAMIDIS, S. & EL-KASSABY, Y.A. 2015. Machine learning use in predicting interior spruce wood density utilizing progeny test information. *Neural Computing and Applications* **28**(3): 505–519. Springer London.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DMINIC, I., KOZINA, A., BAZOK, R. & BARCIC, J.I. 2010. Geographic information systems (GIS) and entomological research: A review. *Journal of Food, Agriculture & Environment* **8**(2): 1193–1198.

DRAKE, V.A. 1994. The influence of weather and climate on agriculturally important insects: an Australian perspective. *Australian Journal of Agricultural Research* **45**: 487.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

EKESI, S. 2010. Combating Fruit Flies in Eastern and Southern Africa (COFESA): Elements of a Strategy and Action Plan for a Regional Cooperation Program. Washington DC.

EKESI, S. & BILLAH, M.K. 2007. *A field guide to the management of economically important Tephritid fruit flies in Africa*. ICIPE Science press, Kenya.

EKESI, S., LUX, S.A. & BILLAH, M. 2005. Field comparison of food-based synthetic attractants and traps for African Tephritid fruit flies. *Development of Improved Attractants and Their Integration into Fruit Fly SIT Management Programmes*. 205–230. International Atomic Energy Agency, Vienna.

EKESI, S., NDERITU, P.W. & RWOMUSHANA, I. 2006. Field infestation, life history and demographic parameters of the fruit fly *Bactrocera invadens* (Diptera: Tephritidae) in Africa. *Bulletin of Entomological Research* **96**: 379–386.

ELIOPOULOS, P.A. 2007. Evaluation of commercial traps of various designs for capturing the olive fruit fly *Bactrocera oleae* (Diptera: Tephritidae). *International Journal of Pest Management* **53**(3): 245–252.

ELITH, J. & LEATHWICK, J.R. 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics* **40**: 677–697.

ENKERLIN, W., GUTIÉRREZ-RUELAS, J.M., CORTES, A.V., ROLDAN, E.C., MIDGARDEN, D., LIRA, E., LÓPEZ, J.L.Z., HENDRICHS, J., LIEDO, P. & ARRIAGA, F.J.T. 2015. Area Freedom in Mexico from Mediterranean Fruit Fly (Diptera: Tephritidae): A Review of Over 30 Years of a Successful Containment Program Using an Integrated Area-Wide SIT Approach. *Florida Entomologist* **98**(2): 665–681.

ENKERLIN, W., VILLASEÑOR, A., FLORES, S., MIDGARDEN, D., LIRA, E., RENDON, P., HURLEY, J., SALAZAR, E., MENDEZ, W., CASTAÑEDA, R., COTOC, E., ZAVALA, J.L., CELEDONIO, H. & GUTIERREZ-RUELAS, J.M. 2016. Descriptive analysis of the factors affecting the population dynamics of the Mediterranean fruit fly (*Ceratitis capitata*, Wied.) in coffee areas located in Guatemala and its implications in IPM Strategies. In: Sabater-Muñoz, B., Vera, M.T., Pereira, R. & Orankanok, W. (Eds.) *Proceedings IX International Symposium on Fruit Flies of Economic Importance, 12-14 May 2014*. 46–63.

EPSKY, N.D., ESPINOZA, H.R., KENDRA, P.E., ABERNATHY, R., MIDGARDEN, D. & HEATH, R.R. 2010. Effective Sampling Range of a Synthetic Protein-Based Attractant for *Ceratitis capitata* (Diptera: Tephritidae). *Journal of Economic Entomology* **103**(5): 1886–1895.

FLETCHER, B.S. 1989. Movements of Tephritid fruit flies. In: Robinson, A.S. & Hooper, G. (Eds.) *World Crop Pests: Fruit Flies - Their Biology, Natural Enemies and Control*. 209–219 (3B edition). Elsevier Amsterdam, Amsterdam.

FRUITFLY AFRICA. 2017. FruitFly Africa.

GETIS, A. & ORD, J.K. 1992. The Analysis of Spatial Association. *Geographical Analysis* **24**(3): 189–206.

GETIS, A. & ORD, J.K. 1996. Local spatial statistics: an overview. In: Longley, P. & Batty, M. (Eds.) *Spatial analysis: modeling in a GIS environment*. 261–277. Geoinformation International, Cambridge, United Kingdom.

GILMORE, J.E. 1989. Sterile Insect Technique (SIT). In: Robinson, A.S. & Hooper, G. (Eds.) *World Crop Pests: Fruit Flies - Their Biology, Natural Enemies and Control*. 353–363 (3B edition). Elsevier Amsterdam, Amsterdam.

GUIDOTTI, D., RAGAGLINI, G. & PETACCHI, R. 2005. Analysis of spatio-temporal *Bactrocera oleae* (Diptera, Tephritidae) infestation distributions obtained from a large-scale monitoring network and its importance to IPM. *Integrated Protection of Olive Crops IOBC/wprs Bulletin* **28**(9): 13–18.

GUISAN, A. & ZIMMERMANN, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.

HAMET, P. & TREMBLAY, J. 2017. Artificial intelligence in medicine. *Metabolism: Clinical and Experimental* **69**: S36–S40. Elsevier Inc.

HANIOTAKIS, G.E. 1974. Sexual Attraction in the olive fruit fly, *Dacus oleae* (Gmelin). *Environmental Entomology* **3**(June): 82–86.

HENDRICHS, J., REYES, J., VREYSEN, M.J.B. & CARDOSO-PEREIRA, R. 2015. Area-Wide Management of Major Pest Insects. In: Sabater-Munoz, B., Pena, L., Navarro, L. & Moreno, P. (Eds.) *XII International Citrus Congress - International Society of Citriculture*. 67–72. Acta Horticulture, Valencia (Spain).

HENDRICHS, M.A., WORNOAYPORN, V., KATSOYANNOS, B.I. & HENDRICHS, J. 2007. Quality Control Method To Measure Predator Evasion in Wild and Mass-Reared Mediterranean Fruit Flies (Diptera: Tephritidae). *Florida Entomologist* **90**(1): 64–70.

HORTGRO. 2016. Key Deciduous Fruit Statistics 2016. Stellenbosch.

JANG, E.B., KLUNGNESS, L.M. & MCQUATE, G.T. 2007. Extension of the use of Augmentoria for Sanitation in a cropping system susceptible to the alien terphritid fruit flies (Diptera: terphritidae) in Hawaii. *Journal of Applied Sciences and Environmental Management* **11**(2): 239–248.

JORDAN, M.I. & MITCHELL, T.M. 2015. Machine learning: Trends, perspectives, and prospects. *Science* **349**(6245): 255–60.

JOTHI, N., RASHID, N.A. & HUSAIN, W. 2015. Data Mining in Healthcare - A Review. *Procedia Computer Science* **72**: 306–313. Elsevier Masson SAS.

JOUBERT, C.J. 1932. Die vrugtevlieg. *Boerdery in Suid-Afrika* **15**: 801–806.

KAPATOS, E. & FLETCHER, B.S. 1983. Seasonal changes in the efficiency of McPhail traps and a model for estimating olive fly densities from trap catches using temperature data. *Entomologia experimentalis et applicata* **33**(1): 20–26. Wiley Online Library.

KARSTEN, M., BARTON, M., ADDISON, P., ADDISON, M.F. & TERBLANCHE, J.S. 2018. A synthesis for managing invasions and pest risks simultaneously for tephritid fruit flies in South Africa. *Entomologia Experimentalis et Applicata*(2): 1–13.

KATSOYANNOS, B.I., KOULOUSSIS, N.A. & CAREY, J.R. 1998. Seasonal and annual occurrence of Mediterranean fruit flies (Diptera: Tephritidae) on Chios Island, Greece: differences between two neighboring citrus orchards. *Annals of the Entomological Society of America* **91**(1): 43–51.

KITRON, U. 1998. Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. *Journal of Medical Entomology* **35**(4): 435–445.

KLASSEN, W. 2005. Area-wide Integrated Pest Mangement and the Sterile Insect Technique. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 39–68. Springer Netherlands, Dordrecht.

KLUNGNESS, L.M., JANG, E.B., MAU, R.F.L., VARGAS, R.I., SUGANO, J.S. & FUJITANI, E. 2005. New sanitation techniques for controlling Tephritid Fruit Flies (Diptera: Tephritidae) in Hawaii. *Journal of Applied Sciences and Environmental Management* **9**(2): 5–14.

KNIPLING, E.F. 1960. The eradication of the Screw-Worm Fly. *Scientific American* **203**(4): 54–61.

KOGAN, M. 1998. Integrated Pest Management: Historical Perspectives and Contemporary Developments. *Annual Review of Entomology* **43**(1): 243–270.

KOUNATIDIS, I., PAPADOPOULOS, N.T., MAVRAGANI-TSIPIDOU, P., COHEN, Y., TERTIVANIDIS, K., NOMIKOU, M. & NESTEL, D. 2008. Effect of elevation on spatio-temporal patterns of olive fly (*Bactrocera oleae*) populations in northern Greece. *Journal of Applied Entomology* **132**: 722–733.

LIEBHOLD, A.M., ROSSI, R.E. & KEMP, W.P. 1993. Geostatistics and Geographic Information Systems in Applied Insect Ecology. *Annual Review of Entomology* **38**: 303–327.

LINDQUIST, D.A. 2000. Pest Management Strategies: Area-wide and Conventional. In: Tan, K. (Ed.) *Area-wide control of fruit flies and other insect pests. Joint proceedings of the international conference on area-wide control of insect pests and the fifth international symposium on fruit flies of economic importance*. 13–19. Penerbit Universiti Sains Malaysia, Penang.

LYONS, D.B., SANDERS, C.J. & JONES, G.C. 2002. The use of geostatistics and GIS as tools for analyzing pheromone trap data at a landscape level: an update. *IOBC/WPRS Bulletin* **25**: 1–14.

MANRAKHAN, A., VENTER, J.H. & HATTINGH, V. 2015. The progressive invasion of *Bactrocera dorsalis* (Diptera: Tephritidae) in South Africa. *Biological Invasions* **17**(7): 2803–2809. Springer International Publishing.

DE MEYER, M., ROBERTSON, M.P., PETERSON, A.T. & MANSELL, M.W. 2008. Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). *Journal of Biogeography* **35**(2): 270–281.

MIDGARDEN, D. & LIRA, E. 2006. Ecological relationship of Medfly and coffee in guatemala and Mexico. *Fruit Flies of Economic Importance: From Basic to Applied Knowledge Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*: 241–247.

MIDGARDEN, D.G., LIRA, E. & SILVER, M. 2014. Spatial Analysis of Tephritid Fruit Fly Traps. In: Shelly, T., Epsky, N., Jang, E.B., Reyes-Floris, J. & Vargas, R. (Eds.) *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. 277–320. Springer, Dordrecht.

MIDGARDEN, D.G., YOUNGMAN, R.R. & FLEISCHER, S.J. 1993. Spatial Analysis of Counts of Western Corn Rootworm (Coleoptera, Chrysomelidae) Adults on Yellow Sticky Traps in Corn - Geostatistics and Dispersion Indices. *Environmental Entomology* **22**(5): 1124–1133.

MITCHELL, T.M. 2006. The discipline of machine learning. Carnegie Mellon University, School of Computer Science, Machine Learning Department.

MORAN, P.A.P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B* **10**: 243–251.

MUMFORD, J. & TWEEN, G. 1997. Economic feasibility study for the control of the Mediterranean and Natal fruit fly in the Western Cape Province. Expert Mission Report RU-7135. Vienna.

MWATAWALA, M.W., DE MEYER, M., MAKUNDI, R.H. & MAERERE, A.P. 2006. Seasonality and host utilization of the invasive fruit fly, *Bactrocera invadens* (Dipt., Tephritidae) in Central Tanzania. *Journal of Applied Entomology* **130**(9–10): 530–537.

MYBURGH, A.C. 1962. Mating habits of the fruit flies *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.). *South African Journal of Agricultral Science* **5**(3): 457–464.

NAVARRO-LLOPIS, V., ALFARO, F., DOMÍNGUEZ, J., SANCHIS, J. & PRIMO, J. 2008. Evaluation of traps and lures for mass trapping of Mediterranean fruit fly in citrus groves. *Journal of economic entomology* **101**(1): 126–131.

NESTEL, D., KATSOYANNOS, B.I., NEMNY-LAVY, E., MENDEL, Z. & PAPADOPOULOS, N.T. 2002. Spatial analysis of Medfly populations in heterogeneous landscapes. *Proceedings of 6th International Fruit Fly Symposium*. 35–43. Stellenbosch, South Africa.

NEUENSCHWANDER, P. & MICHELAKIS, S. 1979. McPhail trap captures of *Dacus oleae* (Gmel.)(Diptera, Tephritidae) in comparison to the fly density and population composition as assessed by sondage technique in Crete, Greece. *Mitteilungen der Schweizerischen Entomologischen Gesellschaft* **52**(4): 343–357.

NYAMUKONDIWA, C. & TERBLANCHE, J.S. 2009. Thermal tolerance in adult Mediterranean and Natal fruit flies (*Ceratitis capitata* and *Ceratitis rosa*): Effects of age, gender and feeding status. *Journal of Thermal Biology* **34**(8): 406–414. Elsevier.

NYAMUKONDIWA, C. & TERBLANCHE, J.S. 2010. Within-generation variation of critical thermal limits in adult Mediterranean and Natal fruit flies *Ceratitis capitata* and *Ceratitis rosa*: Thermal history affects short-term responses to temperature. *Physiological Entomology* **35**(3): 255–264.

OLDEN, J.D. & JACKSON, D.A. 2002. A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* **47**(10): 1976–1995.

OLDEN, J.D., LAWLER, J.J. & POFF, N.L. 2008. Machine Learning Methods Without Tears: A Primer for Ecologists. *The Quarterly Review of Biology* **83**(2): 171–193.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I., KOULOUSSIS, N.A., HENDRICHS, J., CAREY, J.R. & HEATH, R.R. 2001. Early detection and population monitoring of *Ceratitis capitata* (Diptera: Tephritidae) in a mixed-fruit orchard in northern Greece. *Journal of economic entomology* **94**(4): 971–978.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PEÑARRUBIA-MARÍA, E., AVILLA, J. & ESCUDERO-COLOMAR, L.A. 2012. Survival of wild adults of *Ceratitis capitata* (Wiedemann) under natural winter conditions in North East Spain. *Psyche* **2012**: 1–6.

PIMENTEL, R., LOPES, D.J.H., MEXIA, A.M.M. & MUMFORD, J.D. 2014. Spatial regression analysis of *Ceratitis capitata* Wiedemann (Diptera: Tephritidae) on Terceira Island, Azores. *International Journal of Pest Management* **60**(3): 217–223.

PLANT, R.E. & CUNNINGHAM, R.T. 1991. Analyses of the Dispersal of Sterile Mediterranean Fruit-Flies (Diptera, Tephritidae) Released from a Point-Source. *Environmental Entomology* **20**(6): 1493–1503.

PROKOPY, R.J., MILLER, N.W., PIÑERO, J.C., BARRY, J.D., TRAN, L.C., ORIDE, L. & VARGAS, R.I. 2003. Effectiveness of GF-120 fruit fly bait spray applied to border area plants for control of melon flies (Diptera: Tephritidae). *Journal of Economic Entomology* **96**(5): 1485–1493.

PROKOPY, R.J., ROITBERG, B.D. & VARGAS, R.I. 1994. Effects of egg load on finding and acceptance of host fruits in *Ceratitis capitata* flies. *Physiological Entomology* **19**(2): 124–132.

RIBES-DASI, M., ALBAJES, R., SARASÚA, M.J. & AVILLA, J. 2001. The use of geostatistics to study the spatial distribution of *Cydia pomonella* and *Pandemis heparana* in Lleida (Spain). *International Conference on Integrated Fruit Protection* **24**(5): 185–188. International Organization for Biological and Integrated Control, West Palearctic Section, Lleida, Spain.

RIVNAY, E. 1954. The Mediterranean fruit fly, *Ceratitis capitata* Wied.; studies on its toxicants, mass migration, and control in Israel. *Ktavim* **4**: 3–38.

ROBACKER, D.C. 2015. Effects of Shape and Size of Colored Traps on Attractiveness to Irradiated, Laboratory-Strain Mexican Fruit Flies (Diptera: Tephritidae). *Florida Entomologist* **75**(2): 230–241.

ROBACKER, D.C., MORENO, D.S. & WOLFENBARGER, D.A. 1990. Effects of trap color, height, and placement around trees on capture of Mexican fruit flies (Diptera: Tephritidae). *Journal of Economic Entomology* **82**(2): 412–419.

ROESSLER, Y. 1989. Insecticidal bait and cover sprays. In: Robinson, A.S. & Hooper, G. (Eds.) *World Crop Pests: Fruit Flies - Their Biology, Natural Enemies and Control*. 329–335. Elsevier, Amsterdam.

RÖSSLER, Y. 1989. Insecticidal Bait and Cover Sprays. In: Robinson, A.S. & Hooper, G. (Eds.) *World Crop Pests: Fruit Flies - Their Biology, Natural Enemies and Control*. 329–336 (3B edition). Elsevier Amsterdam, Amsterdam.

SCIARRETTA, A. & TREMATERRA, P. 2006. Geostatistical characterization of the spatial distribution of *Grapholita molesta* and *Anarsia lineatella* males in an agricultural landscape. *Journal of Applied Entomology* **130**(2): 73–83.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

STONEHOUSE, J., AFZAL, M., ZIA, Q., MUMFORD, J., POSWAL, A. & MAHMOOD, R. 2002. 'Single-killing-point' field assessment of bait and lure control of fruit flies (Diptera:Tephritidae) in Pakistan. *Crop Protection* **21**: 651–659.

TREMATERRA, P., GENTILE, P. & SCIARRETTA, A. 2004. Spatial analysis of pheromone trap catches of codling moth (*Cydia pomonella*) in two heterogeneous agro-ecosystems, using geostatistical techniques. *Phytoparasitica* **32**(4): 325–341.

VALE, G.A. & PHELPS, R.J. 1978. Sampling problems with Tsetse Flies (Diptera: Glossinidae). *Journal of Applied Ecology* **15**(3): 715–726.

VARGAS, R.I., PIÑERO, J.C., MAU, R.F.L., JANG, E.B., KLUNGNESS, L.M., MCINNIS, D.O., HARRIS, E.B., MCQUATE, G.T., BAUTISTA, R.C. & WONG, L. 2010. Area-wide suppression of the Mediterranean fruit fly, Ceratitis capitata, and the Oriental fruit fly, Bactrocera dorsalis, in Kamuela, Hawaii. *Journal of Insect Science* **10**(135): 135.

VARGAS, R.L., HARRIS, E.J. & NISHIDA, T. 1983. Distribution and Seasonal Occurrence of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) on the Island of Kauai in the Hawaiian Islands. *Environmental Entomology* **12**: 303–310.

VAYSSIÈRES, J.F., KORIE, S., COULIBALY, O., TEMPLE, L. & BOUEYI, S.P. 2008. The mango tree in northern Benin: cultivar inventory, yield assessment, infested stages and loss due to fruit flies (Diptera Tephritidae). *Fruits* **63**(6): 1–8.

VAYSSIÈRES, J.F., SINZOGAN, A., KORIE, S., OUAGOUSSOUNON, I. & THOMAS-ODJO, A. 2009. Effectiveness of Spinosad Bait Sprays (GF-120) in Controlling Mango-Infesting Fruit Flies (Diptera: Tephritidae) in Benin. *Journal of Economic Entomology* **102**(2): 515–521.

VREYSEN, M.J.B. 2000. *Glossina austeni* (Diptera: Glossinidae) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. *Journal of Economic Entomology* **93**(1): 123–135.

VREYSEN, M.J.B. 2005. Monitoring Sterile and Wild Insects in Area-wide Integrated Pest Management Programmes. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 325–361. Springer Netherlands, Dordrecht.

VREYSEN, M.J.B. & SALEH, K.M. 2001. Long-term sampling of gamma sterilised male *Glossina austeni* (Diptera: Glossinidae) with sticky panels on Unguja Island (Zanzibar). *Acta Tropica* **80**(1): 29–37.

WHITE, I.M. & ELSON-HARRIS, M.M. 1992. *Fruit flies of economic significance: their identification and bionomics*. CAB International/Australian Centre for International Agricultural Research, Wallingford, UK/Canberra.

WYSS, J.H. 2000. Screwworm eradication in the Americas. *Annals of the New York Academy of Sciences* **916**: 186–193.

Chapter 2

Spatial patterns and associations of *Ceratitis capitata* and *Ceratitis quilicii* (Diptera: Tephritidae) females in a small-scale heterogeneous landscape in the Western Cape, South Africa

Abstract

Globally and in South Africa, tephritid fruit flies are considered major pests of fruit and vegetable crops. In the Western Cape, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) is the most economically important fruit fly in fruit orchards, causing fruit damage and posing a phytosanitary risk, followed by the recently described *Ceratitis quilicii* De Meyer, Mwatawala and Virgilio (Diptera: Tephritidae). Females lay their eggs underneath the skin of the fruit, causing fruit damage. Understanding the within orchard spatial and temporal distribution and dynamics of these flies, can improve fruit fly management and control. Therefore, the aim of this study was to investigate the spatial distribution and spatial associations between *C. capitata* and *C. quilicii* female trap catch in a heterogeneous orchard environment, at the Welgevallen experimental farm, Stellenbosch, South Africa. Weekly adult female trap catch data was collected for the 2016/2017 fruiting season, from 70 biolure fruit fly traps, evenly spaced in and around the Welgevallen orchard, this include commercial fruit crops, home gardens, natural vegetation and a pine hedge. Weekly female trap catch were used to determine the spatial distribution patterns of *C. capitata* and *C. quilicii* as well as the spatial association between the two species, using spatial analysis by distance indices (SADIE). *Ceratitis capitata* was the dominant fly at Welgevallen in all habitat types, except for the natural vegetation, nectarines and the pine hedge where the proportion of *C. quilicii* was higher. Both species had aggregated spatial patterns, however, *C. capitata* significantly aggregated more towards the end of the season while *C. quilicii* significantly aggregated at the beginning of the season. *Ceratitis capitata* and *C. quilicii* females were

spatially associated, most prominently in home gardens, natural vegetation, citrus and nectarines. The results also suggest dispersal of the two species from unmanaged home gardens and natural vegetation to commercial orchards, based on the aggregation of these flies in these habitats as well as on the edge of the commercial fruit orchard, adjacent or near to these habitats. This information is valuable to incorporate into the decision-making processes of an integrated pest management strategy.

Introduction

Tephritid fruit flies are known to cause major damage to fruit and vegetable crops around the world, while many also hold a phytosanitary risk for countries, which export fruit and vegetable crops (White & Elson-Harris 1992). In South Africa, which has a diverse climate and agricultural landscape, numerous tephritid fruit fly species are considered pests of cultivated plants (see Karsten *et al.* 2018). In the Western Cape, the agroecosystem consists of a heterogeneous landscape, comprising large pome fruit, stone fruit, citrus and grape (wine grapes and table grapes) production areas. *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae), commonly known as the Mediterranean fruit fly or Medfly, is the most dominant and economically important fruit fly occurring in this region (DeVilliers *et al.* 2013). Another economically important fruit fly, but less widely distributed in the Western Cape (DeVilliers *et al.* 2013) is the newly described species, *Ceratitis quilicii* De Meyer, Mwatawala and Virgilio (Diptera: Tephritidae) (De Meyer *et al.* 2016), which belongs to the *Ceratitis* FAR complex (Barr & McPheron 2006; Virgilio *et al.* 2008). This species was previously thought to be a *Ceratitis rosa* Karsch morphotype, referred to as “R2”, “the cold type” or “highland type” in previous literature (see De Meyer *et al.* 2015; Hendrichs *et al.* 2015, also see the references therein), whereas the other *C. rosa* morphotype was previously referred to as “R1”, “the hot type”, or “lowland type”. However, a review of these morphotypes by De Meyer *et al.* (2016), classified these morphotypes into two separate species belonging to the *Ceratitis* FAR complex, which consist of a number of species including *Ceratitis fasciventris* (Bezzi), *Ceratitis anonae* Graham, *C. rosa* Karsch (Barr & McPheron 2006; Virgilio *et al.* 2008) and recently *C. quilicii*. Since the description of *C. quilicii,* no record of *C. rosa* from the Western Cape has been published. Therefore, it can be assumed that all previous research conducted on ‘*C. rosa’* in the Western Cape (or research conducted where flies originated from the Western Cape) were on *C. quilicii*. However, this would need to be confirmed with new trapping records and the revision of all museum specimens originating

from the Western Cape, as no large scale surveys have taken place since the renaming of the species two years ago.

The fact that both *C. capitata* and *C. quilicii* are economically important fruit flies makes it important to manage these flies effectively, as part of any integrated pest management programme. Understanding the within orchard spatio-temporal distribution and dynamics of these flies, can improve the location and timing of management methods and also highlight the areas where management is not effective (Papadopoulos *et al.* 2003). The spatio-temporal distribution of *C. capitata* populations in fruit orchards have been studied widely (see Israely *et al.* 2005a, 2005b; Nestel *et al.* 2002; Papadopoulos *et al.* 2003; Sciarretta and Trematerra 2011), mostly in Europe and Israel. *Ceratitis capitata* has been found to respond to its variable environment and exploits areas where the environment favours some aspects of the flies' biology and ecology, regarding suitable climate, hosts, food, mates or refuge (Nestel *et al.* 2002). Host fruit and ripening sequence, but also human interference in terms of application of insecticides and other fruit fly control methods, have been found to influence the changes in spatial distributions of *C. capitata* in fruit orchards (Sciarretta & Trematerra 2011). Unmanaged hosts near commercial orchards and in home gardens, play an important role in the spatio-temporal distribution of *C. capitata* and *C. quilicii.* They act as refuges and breeding grounds during times when intensive control is being applied in commercial orchards (Israely *et al.* 1997) and when no commercial fruit are available during winter (Myburgh 1956; DeVilliers *et al.* 2013). Flies disperse from these refuges and breeding grounds as soon as commercial fruit become available, and susceptible to attack (Myburgh 1956). It is well known that tropical fruit flies are strong flyers, which will disperse in search of food, shelter, oviposition sites and mating partners (Bateman 1972).

Although extensive spatial analysis has been conducted to determine the spatial distribution and dynamics of *C. capitata* in fruit orchards in Europe and Israel, no study has looked at the spatial patterns and spatial association of *C. capitata* and *C. quilicii* in the Western Cape, South Africa. Therefore, the aim of this study was to investigate the spatial patterns and spatial associations between *C. capitata* and *C. quilicii* trap catch, making use of a high-density trapping grid in a mixed fruit orchard. The first objective was to quantify the spatial patterns of each species at different times of the season, in terms of regular, random, or aggregated spatial patterns. Furthermore, to determine the association between the spatial patterns of these two economically important fruit fly pests at different times of the season. The results are discussed in the context of the development and implementation of integrated pest management (IPM) strategies against *C. capitata* and *C. quilicii* in mixed fruit orchards with varying fruit phenologies in the Western Cape, South Africa, and will contribute new information on host status and occurrence of *C. quilicii*.

Materials and methods

Study area

The study was conducted at the Welgevallen experimental farm (-33.947822°, 18.871680°) on the Stellenbosch University campus, Stellenbosch, Western Cape, South Africa. The farm comprises of mixed fruit agriculture, while the orchard area (4.23 ha) used for this study consisted of citrus (Rutaceae), variety Satsuma (1.34 ha), plums (Rosaceae), variety Laetitia and Songold (0.62 ha), apples (Rosaceae), variety Royal Gala and Pink Lady (0.51 ha), pears (Rosaceae), variety Forelle (1.1 ha) and nectarines (Rosaceae), variety Alpine (0.66 ha). The area surrounding the orchard comprised of home gardens containing *Dovyalis caffra* (J.D. Hook & Harvey) J.D. Hook, commonly known as kei-apple, *Eriobotrya japonica* (Thunb.) Lindley, commonly known as loquat, *Psidium guajava* L. (guava) as well as *Pyrus spp.* and *Prunus* spp. Furthermore, native species, mostly wild olive trees *Olea europaea*

subsp. Africana (Mill.) P.S. Green and *Podocarpus* spp., but also timber plantations of *Eucalyptus* spp., were found in close proximity of the orchard. For the purpose of this chapter, the *O. europaea subsp. Africana*, *Podocarpus* spp. and *Eucalyptus* spp. were classified as natural vegetation to distinguish between the orchard, home gardens and the other surrounding habitats. Other land uses near the orchard area included pastures.

Weather data for the study area was obtained from a weather station (-33.959088°, 18.833699°; 125 m elevation), situated approximately 3.5 km from the study site, which is operated by the Agricultural Research Council (ARC) of South Africa.

Adult fruit fly control

A series of different products and application methods were applied against *Ceratitis* spp. during the study period, including the Bait Application Technique (BAT) and cover sprays. Generally, the products were applied to the orchard area on a weekly basis depending on weather conditions. A summary of the application detail of control methods are provided in Table 1.

Table 1. A summary of the pest control measures in the mixed fruit orchard on Welgevallen experimental farm.

Time (weeks)	Year	Type of application	Product	Active ingredient	Target species	Orchard crops
33–42	2016	Baiting	HYM-LURE + MERCAPTOTHION	Protein hydrolysate 435 g/l + Mercaptothion 500 g/l	*Ceratitis* spp.	All
43	2016	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	All
44	2016	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	Nectarines
	2016	Baiting	HYM-LURE + MERCAPTOTHION	Protein hydrolysate 435 g/l + Mercaptothion 500 g/l	*Ceratitis* spp.	All except nectarines
45 -48	2016	Cover spray	AZINPHOS 200SC	Azinophos-methyl 200 g/l	*Grapholita molesta* Busck (Lepidoptera: Tortricidae)	Plums
45	2016	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	All
46	2016	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	Nectarines
46	2016	Baiting	HYM-LURE + MERCAPTOTHION	Protein hydrolysate 435 g/l + Mercaptothion 500 g/l	*Ceratitis* spp.	All except nectarines
47	2016	Cover spray	CYPERIN	Cypermethrin 200 g/l	*Ceratitis* spp.	Nectarines
47	2016	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	All except nectarines
48 – 51	2016	Baiting	HYM-LURE + MERCAPTOTHION	Protein hydrolysate 435 g/l + Mercaptothion 500 g/l	*Ceratitis* spp.	All
52 – 16	2016-2017	Baiting	GF-120* NF	Spinosad (naturalyte) 0,24 g/l	*Ceratitis* spp.	All
1 – 5	2017	Cover spray	CALYPSO® 480 SC	Thiacloprid 480 g/l	*Cydia pomonella* Linnaeus (Lepidoptera: Tortricidae)	Apples and pears
	2017	Cover spray	CRYPTOGRANTM	*Cryptophlebia leucotreta* granulovirus (CrleGV-SA) (at least 5 x 1010 OBs/ml)	*Thaumatotibia leucotreta* (Lepidoptera: Tortricidae)	Citrus
17 – 19	2017	Baiting	HYM-LURE + MERCAPTOTHION	Protein hydrolysate 435 g/l + Mercaptothion 500 g/l	*Ceratitis* spp.	All

Adult activity

Weekly fruit fly monitoring was conducted from 19 August 2016 to 13 June 2017. Seventy traps were placed in and around the Welgevallen orchard (Figure 1). Traps placed inside the orchard were spaced 30 m apart. Traps in natural vegetation were placed within a 30–60 m buffer area surrounding the orchard area. In home gardens, which were in close proximity to the orchard (± 30 – 160 m), the traps were placed randomly in home garden fruit fly host trees. Traps in the pine hedge, as in the orchard, were also placed ± 30 m apart.

Figure 1. Location of fruit fly traps at Welgevallen, Stellenbosch, South Africa. Note that the aerial photograph serving as the base of this map was taken in 2014 and does not necessarily depict the crops planted in the 2016/2017 season, however all crops that were planted at the time of this study were monitored. The crops surrounding the orchards were wine grape vineyards, located to the east and south-west as well as grassy fields used for cattle grazing.

Yellow bucket traps baited with Chempac Uni-pack Lure (Biolure) (Chempac Pty Ltd., Paarl, South Africa) were used for monitoring adult fruit flies, targeting female *Ceratitis* spp., the

dominant flies in the Stellenbosch region (DeVilliers *et al.* 2013; Manrakhan & Addison 2014). Every week traps were emptied and the content identified in the laboratory at Stellenbosch University. *Ceratitis quilicii* males were identified based on leg feathering, according to De Meyer *et al.* (2016). Given the fact that *C. rosa* and *C. quilicii* females cannot currently be morphologically differentiated (De Meyer *et al.* 2016), and the fact that all males were *C. quilicii* it was assumed that the female flies that were caught were also *C. quilicii* and not *C. rosa.*

Because female biased traps were used for trapping, the focus of the analysis was on adult female fruit flies. Weekly *C. capitata* and *C. quilicii* female trap counts were recorded as the mean number of flies per trap per week (FTW) per habitat, including each crop type in the orchard, home gardens, pine hedge and natural vegetation.

Data analysis

To assess the differences in the mean female FTW between *C. capitata* and *C. quilicii* in the different habitats, a two-way mixed model analysis of variance (ANOVA) was conducted with habitat type and species as the main effects and traps nested in habitat type as the random effect. Traps were treated as random effects because traps were regarded as a random selection of traps from a larger population. Furthermore, a mixed model ANOVA was selected to take into account the dependencies of the measurements within traps. An LSD Fisher's post-hoc test was conducted in the case of significant group differences. Before analysis, trap counts were transformed to Log10(x + 1) to homogenise variances in the data, while untransformed means (FTW) were used to draw graphs for easier interpretation. Data analysis were conducted using the STATISTICA version 13.3 software package (TIBCO Software Inc 2017).

Fruit damage assessments

Fruit damage assessments were conducted per habitat type. Twenty-five evenly spaced trees were chosen for assessing fruit fly damage in each habitat type within the orchard. This method was adapted from Brown & Pringle (2006) as the standard monitoring system used for pome fruit orchards. On each tree, five randomly selected fruit as well as additional fallen fruit (as available) under the tree were visually inspected for fruit fly damage symptoms, which included oviposition punctures, exit holes and decayed areas on the fruit skins (Manrakhan & Addison 2014). Furthermore, fallen fruit from home garden trees were also collected. No fruits were collected from the pine hedge and the natural vegetation, as no fruits were available in the near vicinity of the traps in these habitats.

Fruit that showed any damage symptoms were collected and placed in plastic bags, keeping the fruit from the tree and the ground separate. The fruit damage assessment of each habitat type within the orchard was conducted within one week of harvest. Fruits in home gardens were collected in week 6 of 2017. The fruit was transported to the laboratory, where the fruit was weighed using a precision balance (RADWAG®, Model PS 4500/c/2) and incubated in five-litre square plastic containers on a layer of 1-2 cm of sterilised vermiculite. The vermiculite was sterilised by freezing it for 48 hours at -40 °C. Incubation was conducted at 25°C with a 12:12 photoperiod, for approximately two months or until no further flies emerged. All reared adult fruit flies were collected and identified, as explained above, sexed and counted. To calculate the degree of fruit infestation the number of fruit flies reared from fruit was divided by the total weight (kg) of the fruit. Fruit infestation (fruit flies per kg fruit) was calculated for fruits sampled from the trees as well as from the ground from each habitat type (different crops in the orchard as well as home gardens).

Spatial analysis

Ceratitis capitata and *C. quilicii* female trap catch were spatially analysed to determine the spatial pattern of each species as well as the spatial association between the trap catch of the two species. Spatial analysis of the trap catch was conducted using SADIE (Spatial Analysis by Distance Indices) (Perry 1995). The SADIE package is used to quantify spatial patterns of organisms using georeferenced count data for a single dataset; however, it can also measure the spatial association between two datasets (e.g. spatial association between two species) (Perry *et al.* 1996). SADIE quantifies spatial patterns either as aggregated, random or regular, by comparing the spatial arrangement of the original data to permuted spatial arrangements, such as the most regular arrangement, derived from the original data. SADIE uses the distance to regularity, *D*, which can be described as the sum of minimum distance moved by each individual from its original spatial arrangement to a position where all individuals are arranged in a regular manner (Perry 1995).

In the case of the fruit fly dataset, *D,* can be described as the sum of the minimum distance which each fruit fly must have travelled so that each trap contained the same number of fruit flies. Using *D,* an overall index of aggregation, I_a, where $I_a = 1$ refers to a random spatial arrangement, $I_a < 1$ refers to a regular spatial arrangement and $I_a > 1$ refers to an aggregated spatial arrangement, were calculated for comparison between different datasets (Perry 1995; Perry *et al.* 1996). SADIE employs the pseudo-random number generator algorithm, AS183, (Wichmann & Hill 1982) in order to conduct probability tests under the null hypothesis of complete randomness at the 5% level. Where $P_a < 0.025$ indicates a statistically significant aggregated spatial arrangement, $P_a > 0.975$ indicates a statistically significant regular spatial arrangement and $0.025 < P_a < 0.975$ indicates a random arrangement (Perry 1998). The degree of clustering of the count data (trap catch) of each sampling unit (trap) within the overall spatial arrangement were also calculated using a

clustering index, v, to compare between different units and between different datasets (Perry *et al.* 1999). Clusters are either patch clusters, which represent a neighbourhood of units with larger values than the sampling mean, or gap clusters, which are units with smaller values than the sampling mean that are close to one another. Units with values exceeding the sampling mean are ascribed a clustering index, v_i, where v_i > 1.5 indicates that the unit belongs to a patch cluster. Similarly, for sampling units with values less than the sampling mean a clustering index, v_j, is ascribed to each unit where v_j < -1.5 indicates that the unit belongs to a gap cluster (Perry *et al.* 1999; Perry & Dixon 2002). To test for overall clustering at the 5 % level, a mean clustering index, $\overline{v}$, was calculated, where $\overline{v}_i$ indicated patches and $\overline{v}_j$ indicated gaps, which were then compared to corresponding values obtained from randomisations by the algorithms in SADIE.

To conduct the spatial pattern analysis in SADIE two parameters were specified, an integer seed, *iseed,* used in the pseudo-random number generator algorithm and an integer value, *k5psim*, which determines the number of randomisations. The *iseed* and *k5psim,* were set to 29 892 and 57 respectively.

Determining whether the two fruit fly species were spatially associated, disassociated, or occurred randomly with respect to one another a local spatial association index, X_k, which is based on the patch and gap clustering indices at the *k*th, sampling unit between the two species was calculated using the SADIE software. Positive X_k values point toward two patches or two gaps from both species that spatially coincide, and indicates spatial association between the species. Negative X_k values point toward a gap of the one species coinciding with a patch from the other species and vice versa, indicating spatial disassociation between the species at the *k*th sampling unit (Perry *et al.* 1999; Perry & Dixon 2002). An overall spatial association index, X, was also calculated between the two species,

testing its significance at the 5 % level through randomisations, by reassigning values among the sampling units, where $P < 0.025$ indicates significant spatial association and $P > 0.975$ indicated significant disassociation. These randomisations are done after allowing for small-scale spatial autocorrelation in the clustering indices of each species using a modified t – Test method developed by Dutilleul *et al.* (1993). As for the spatial pattern analysis, two parameters need to be specified, an integer, *iseed*, used in the pseudo-random number generator algorithm and an integer, *Nsims*, determining the number of randomisations conducted. In this case, the *iseed* and the *Nsims* were specified as 12345 and 10000 respectively.

In terms of spatial patterns the overall spatial aggregation index, I_a, for each trapping week was determined as well as the mean clustering index, $\overline{v}$, for patches and gaps together with their respective probability measures. These analyses were also conducted for the entire sampling period (all weeks) as well as for the first period of sampling – early season (week 35 – 5) and the second period of sampling – late season (week 6 – 25), which was based in fruit fly trapping activity. The spatial association analysis was conducted yielding an overall spatial association index, X, with its corresponding probability measure, *P,* for the same periods.

To visualise the spatial patterns of fruit fly activity from trap counts, the inverse distance weighting interpolated estimates of the pattern and association indices were mapped using ArcGIS 10.5. Spatial pattern maps indicating significant patches ($v_i > 1.5$) and significant gaps ($v_j < -1.5$) for each species were created (Perry & Dixon 2002). Furthermore, *C. capitata* / *C. quilicii* spatial association maps were also created for each period where $X_k > 0.5$ (positive associations) and $X_k < -0.5$ (dissociation).

Results

Adult activity

Ceratitis capitata was the dominant fruit fly species at Welgevallen during the trapping period in most habitat types except for the natural vegetation, nectarines and the pine hedge where the proportion of *C. quilicii* was higher. Traps caught primarily female flies, as was expected (Table 2). Home gardens yielded the most captures (male and female) per trap per week followed by citrus, while the pine hedge traps caught the least. In terms of sex composition of the total captures, traps in apples caught the highest proportion of *C. capitata* females while traps in the natural vegetation caught the lowest proportion of *C. capitata* females. Traps in the pine hedge caught the highest proportion of *C. quilicii* females while traps in citrus caught the lowest (Table 2).

Table 2. The average number of fruit flies caught per trap per week (±SE) and the relative abundance of *Ceratitis capitata* and *Ceratitis quilicii* per habitat type and sex (%) on Welgevallen, Stellenbosch, from 19 August 2016 to 13 June 2017.

		% of total captures			
		C. capitata		*C. quilicii*	
Habitat	**Average captures per trap per week**	**Female**	**Male**	**Female**	**Male**
Apples	1.12 ± 0.20	76.74	5.56	17.36	0.35
Citrus	3.38 ± 0.45	67.45	15.80	13.36	3.39
Home gardens	3.51 ± 0.45	45.57	11.02	35.34	8.07
Natural vegetation	3.30 ± 0.41	18.49	6.13	58.97	16.41
Nectarines	0.83 ± 0.17	40.58	4.71	50.00	4.71
Pears	0.35 ± 0.05	70.86	10.60	16.56	1.99
Pine hedge	0.07 ± 0.02	9.09	9.09	63.64	18.18
Plums	0.71 ± 0.17	45.12	20.12	30.49	4.27

Significant ($F_{(7, 62)}$ = 6,9866, $p < 0.001$) habitat*species interaction was observed in the trapping data. There were significant differences ($p < 0.05$) in the mean number of *C. capitata* and *C. quilicii* female FTW in all habitat types except for home gardens ($p = 0.98$),

the pine hedge (p = 0.16) and nectarines (p = 0.76) (Figure 2). *Ceratitis capitata* females were the dominant fly in all the habitat types where significant differences in the trap counts were observed, except for in the natural vegetation where *C. quilicii* was the dominant fly.

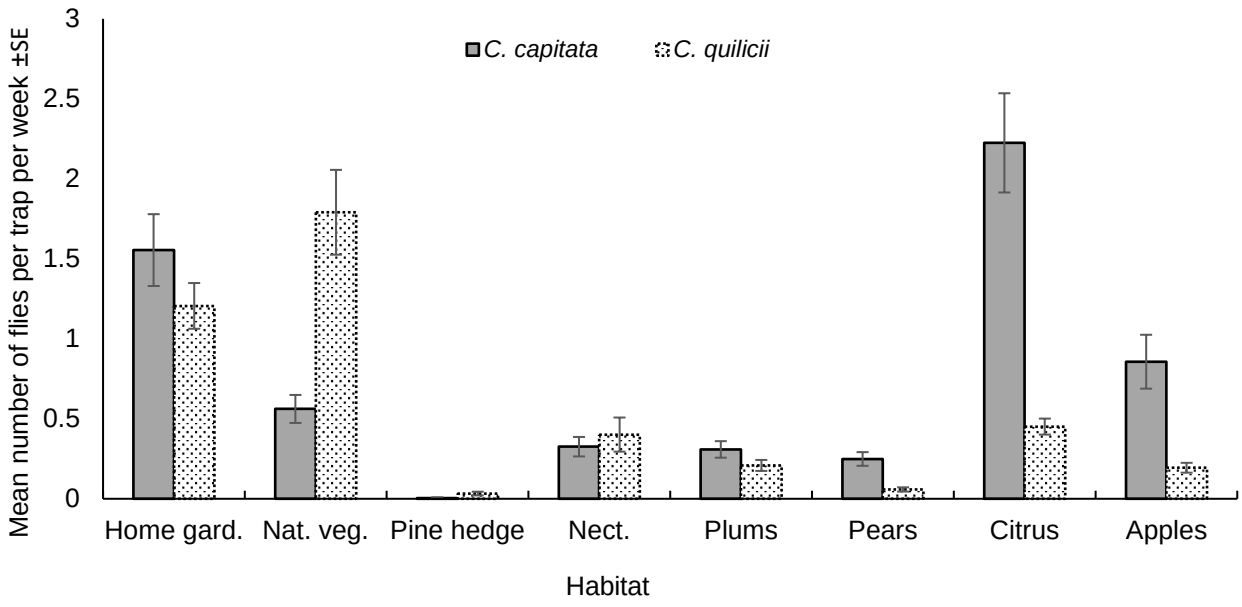

Figure 2. Mean (±SE) *Ceratitis capitata* and *Ceratitis quilicii* females caught per trap per week in each habitat type on Welgevallen, Stellenbosch, South Africa, from 9 August 2016 to 13 June 2017.

C. quilicii females peaked early in the season (between week 45 and 51) after which the population slowly decreased, compared to *C. capitata*, which peaked from week 11 and onwards declining towards week 25 (Figure 3 and 4). However, an exception was observed in the nectarines where an early peak (week 47) in female *C. capitata* trap catches was seen. After week 19, which is the start of winter, both species' trap catches declined to low levels. Generally, *C. capitata* and *C. quilicii* female population peaks occurred during the week of harvest or a few weeks thereafter in the orchard habitats. In home gardens, *C. quilicii* females were the dominant fly early in the season, while *C. capitata* females were caught in greater numbers at the end of the season (Figure 3).

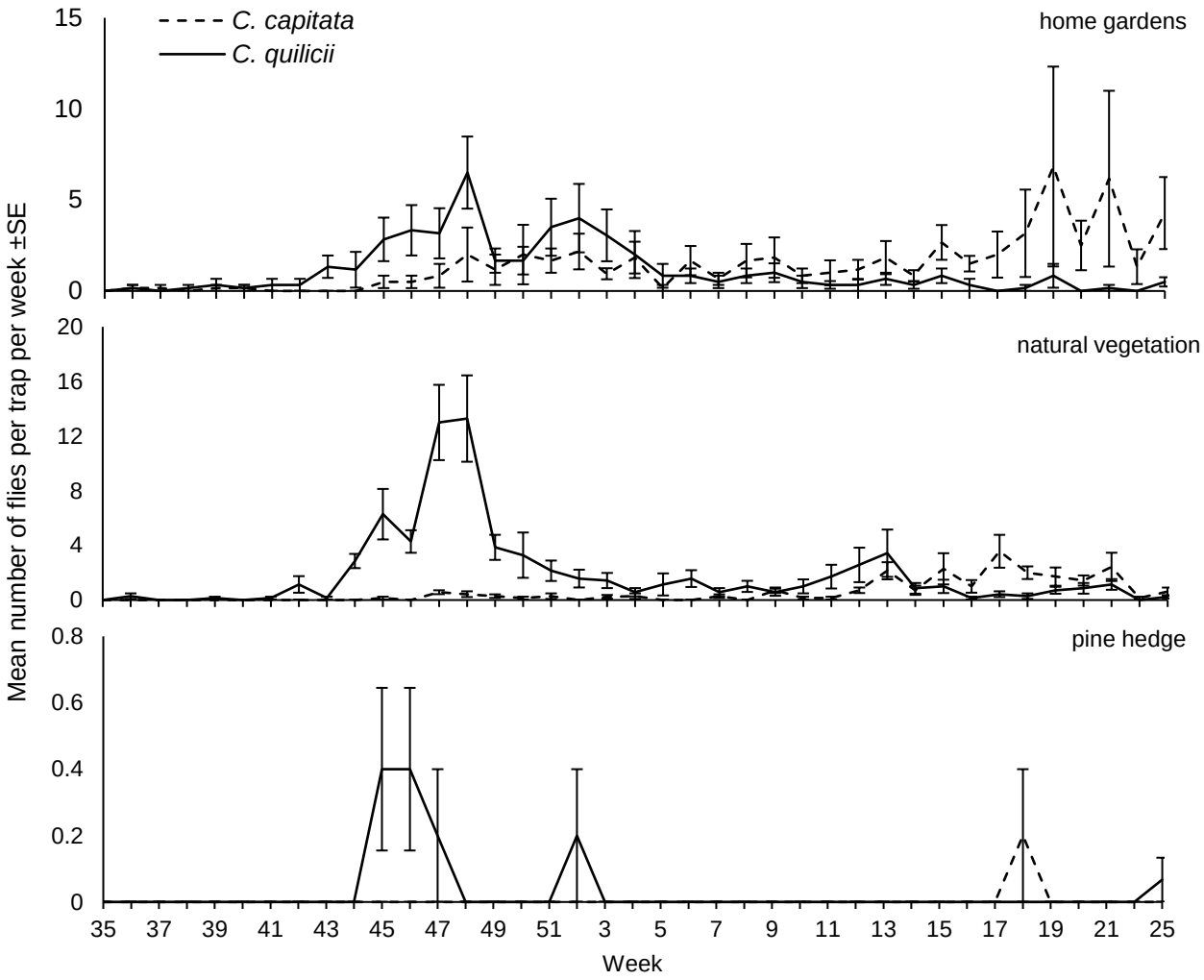

Figure 3. Mean (±SE) *Ceratitis capitata* and *Ceratitis quilicii* females caught per trap per week, in each habitat type, outside of the Welgevallen orchard, Stellenbosch, South Africa.

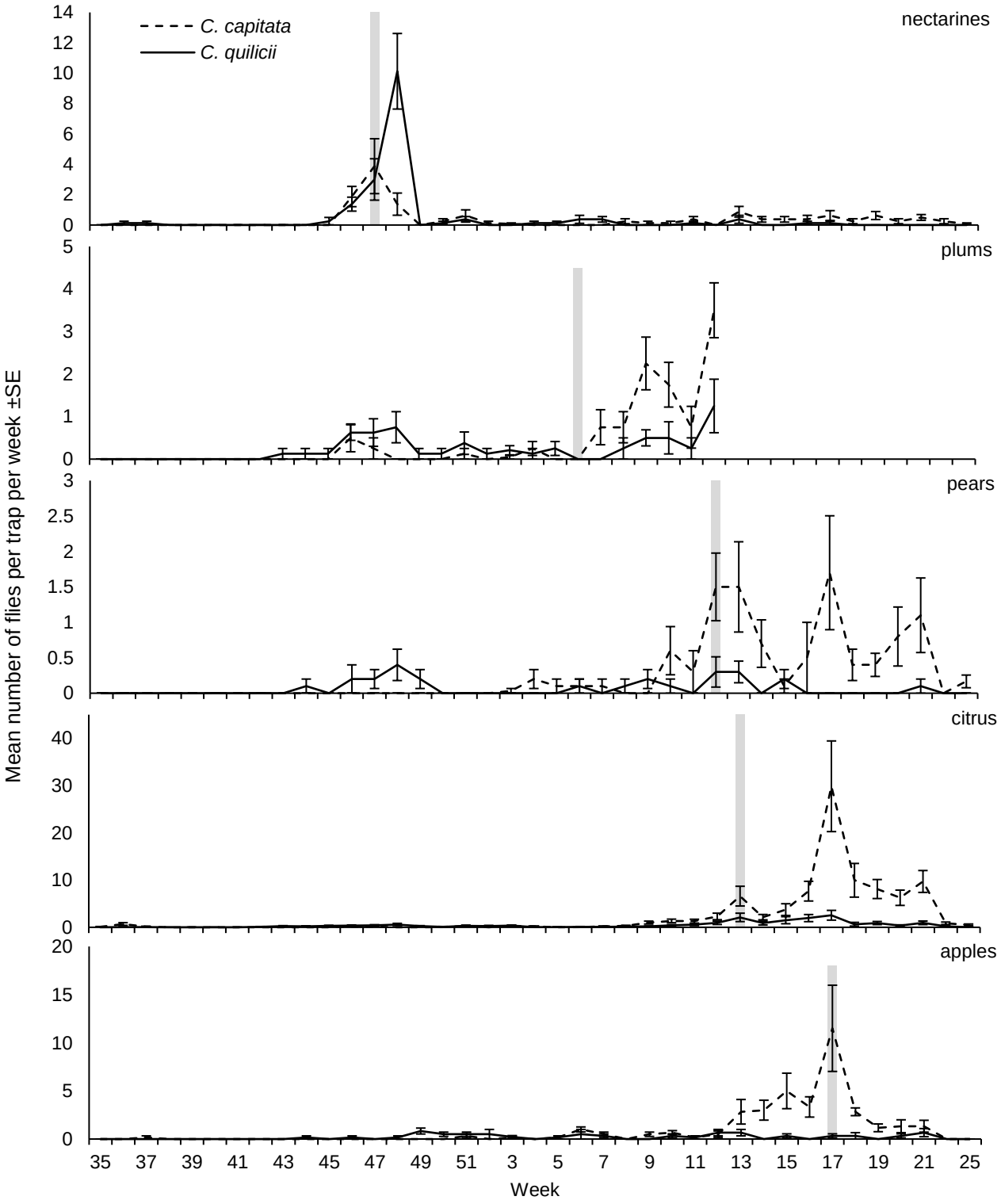

Figure 4. Mean (±SE) *Ceratitis capitata* and *Ceratitis quilicii* females caught per trap per week, in each habitat type, within the Welgevallen orchard, Stellenbosch, South Africa. The grey bars indicate the start of the harvest of the different crops. Note the varying scale on the x-axis to ensure clarity of graph. No trapping data were obtained after week 12 in plums, as the orchard was removed.

Maximum and minimum temperature increased towards their peaks in week 8 and week 6, respectively, after which temperatures dropped. The study area experienced a dry period between week 43 (2016) to week 9 (2017), after which the weekly rainfall started to increase towards the end of the study period (week 25 of 2017). The most rainfall during the study period (week 35 of 2016 to week 25 of 2017) was measured between week 9 and week 25 of 2017 (Figure 5).

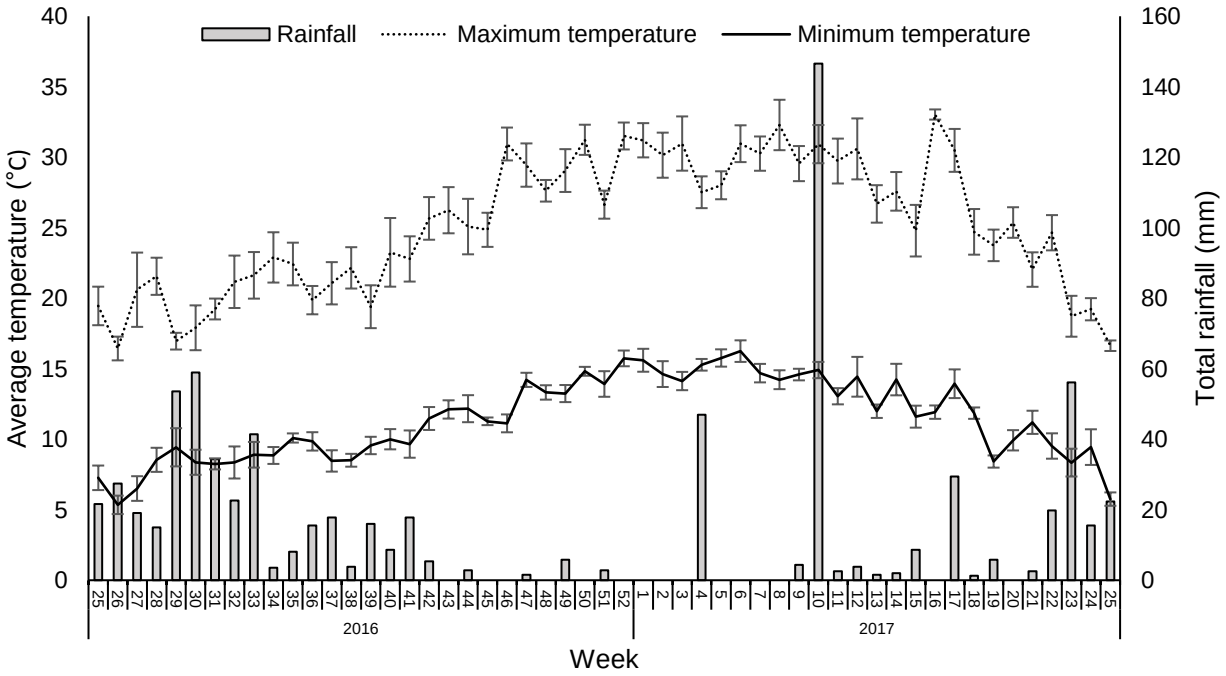

Figure 5. Weekly average minimum and maximum temperature (°C) and total rainfall (mm) for a year including the study period. The error bars indicate standard error.

A list of the tephritid fruit fly by-catch from Biolure bucket traps in the different habitat types in the study area from 9 August 2016 to 13 June 2017 are presented in Table 3. *Ceratitis podocarpi* (Bezzi) (Diptera: Tephritidae) was caught in all habitat types but no flies were reared from fruits. The *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae) individual (male) caught in the natural vegetation was the first record thereof in the Stellenbosch area as well as the second record in the Western Cape, and was trapped on 25 April 2017 in a Biolure bucket trap (-33,946817°, 18,872049°) on the northern border of the fruit orchard in natural vegetation.

Table 3. Total number of tephritid fruit fly by-catch in Biolure bucket traps caught from 9 August 2016 to 13 June 2017 in different habitat types on Welgevallen, Stellenbosch, South Africa.

	Species						
Habitat	*Ceratitis podocarpi* (Bezzi)	*Ceratitis rubivora* (Coquillet)	*Ceratitis aliena* (Bezzi)	*Dacus ciliatus* Loew	*Bactrocera biguttula* (Bezzi)	*Bactrocera oleae* (Rossi)	*Bactrocera dorsalis* (Hendel)
Apples	2	0	0	0	0	0	0
Citrus	60	4	0	2	1	0	0
Home gardens	16	0	0	4	0	0	0
Natural vegetation	50	13	5	2	4	0	1
Nectarines	11	1	0	0	0	0	0
Pears	3	0	0	0	1	1	0
Pine hedge	4	0	0	1	0	0	0
Plums	4	1	0	2	0	0	0

Fruit damage assessments

Ceratitis capitata was the dominant fly reared from all fruits and the only fly reared from apples (Table 4). From the fruits sampled in home gardens, the composition of the two species was similar. Home garden fruits, especially kei-apple were also the most heavily infested with fruit flies, followed by nectarines and citrus, with apples the least infested. In all cases, infestation levels from fruit collected from trees were higher than those collected from the ground. No flies were reared from pears and plums.

Table 4. Fruit infestation (flies/kg/fruit) and % species composition (*Ceratitis capitata* and *Ceratitis quilicii*) from fruit sampled per habitat type at fruit harvest on Welgevallen, Stellenbosch, South Africa.

				% species composition	
Habitat	**Number of flies**	**Weight of fruit (kg)**	**Infestation (flies/kg fruit)**	***C. capitata***	***C. quilicii***
Apples (Trees)	1	0.640	1.56	100	0
Apples (Ground)	2	3.691	0.54	100	0
Citrus (Trees)	12	1.243	9.65	100	0
Citrus (Ground)	11	4.548	2.42	18.2	81.8
Home gardens (Ground)	28	1.084	25.83	53.5	46.4
Natural vegetation	-	-	-	-	-
Nectarines (Trees)	54	4.911	10.99	62.9	37.1
Nectarines (Ground)	2	0.290	6.90	100	0
Pears (Trees)	0	0.304	0	0	0
Pears (Ground)	0	4.903	0	0	0
Pine hedge	-	-	-	-	-
Plums (Trees)	0	4.888	0	0	0
Plums (Ground)	0	8.112	0	0	0
All	110	34.614	3.18	61.8	38.2

Spatial analysis

Ceratitis capitata female trap catches showed an aggregated spatial pattern ($I_a > 1$), for all three periods (all weeks, week 35 to 5 and week 6 to 25), however, none of these spatial patterns were significant ($0.025 < P_a < 0.975$). For the same periods, *C. quilicii* female trap catches also showed an aggregated spatial pattern ($I_a > 1$) with a significant aggregated spatial arrangement for the period, week 35 to week 5 (early season). Significant spatial associations ($P < 0.025$) between *C. capitata* and *C. quilicii* female trap catches were observed during all three periods (Table 5).

Table 5: The spatial patterns of *Ceratitis capitata* and *Ceratitis quilicii* females and the spatial association between the two species during three periods on Welgevallen, Stellenbosch, South Africa.

	Spatial pattern				**Spatial association**	
	Ceratitis capitata		*Ceratitis quilicii*		*Ceratitis capitata / Ceratitis quilicii*	
Period	I_a	P_a	I_a	P_a	X	P
All weeks	1.480	0.0265	1.418	0.0351	**0.5473**	**0.0002**
Week 35 to 5	1.296	0.0873	**1.461**	**0.0238**	**0.2806**	**0.019**
Week 6 to 25	1.465	0.0265	1.266	0.0954	**0.8577**	**0.0001**

I_a = index of aggregation where $I_a = 1$, random spatial pattern; $I_a < 1$, regular spatial pattern and $I_a > 1$, aggregated spatial pattern.
P_a = probability level of aggregation index where $P_a > 0.975$, regular spatial arrangement (underlined); $P_a < 0.025$, aggregated spatial arrangement (bold) and $0.025 < P_a < 0.975$, random spatial arrangement.
X = index of association where $X > 0$, association and $X < 0$ dissociation
P = probability level of association index, where $P < 0.025$, significant association (**bold**)

Ceratitis capitata and *C. quilicii* spatial patterns and spatial associations are visualised in Figure 6. For the entire season, the spatial patterns of *C. capitata* and *C. quilicii* females formed patches, predominantly on the western edge of the orchard and in the home gardens and natural vegetation. Gaps for both species occurred predominantly within the orchard. In the early season and late season, a similar pattern of patches and gaps were seen in both species. *Ceratitis capitata* early season patches occurred at both the western edge (facing the home gardens) of the orchard and in the home gardens and natural vegetation, while later in the season *C. capitata* patches occurred more toward the western edge and within the orchard rather than in the home gardens and natural vegetation. In terms of the early season, a similar patch pattern was observed for *C. quilicii* and *C. capitata*, except for only scattered *C. quilicii* patches on the western edge of the orchard. During the late season, *C. quilicii* and *C. capitata* had similar spatial patterns. There was a strong positive association between the spatial patterns of *C. capitata* and *C. quilicii* during all three periods, predominantly in the home gardens, citrus, nectarines and natural vegetation (Figure 6).

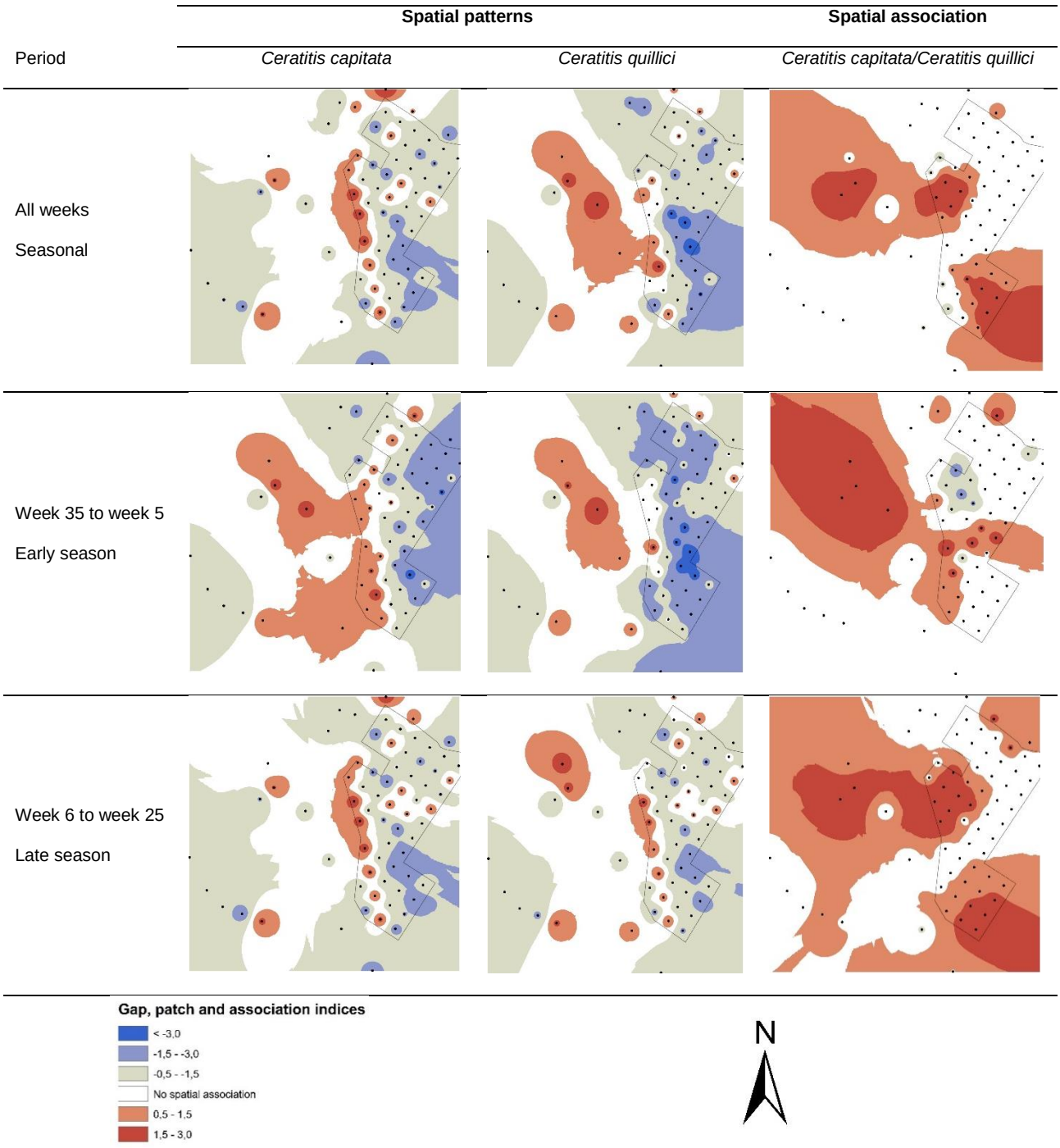

Figure 6. Interpolated spatial clustering and association indices of *Ceratitis capitata* and *Ceratitis quillici* trapping activity in a mixed fruit orchard on Welgevallen, Stellenbosch, South Africa during the 2016/2017 fruiting season. Mapped indices include v_i > 1,5 (patches) and v_j < -1,5 (gaps) for spatial clustering; and X_k > 0.5 (positive association) and X_k < - 0.5 (disassociation) for spatial association. Red indicates patches and association. Blue indicates gaps and disassociation. The darker the colour hue, the stronger the association.

Generally, during the weeks of the early season the spatial patterns of *C. capitata* and *C. quilicii* females were aggregated (I_a > 1). *C. quilicii* females for the first time showed a significant aggregated spatial arrangement in week 43 (P_a < 0.025), while *C. capitata* females only showed significance (P_a < 0.025) in their aggregated spatial arrangements for

the first time in week 52 (Table 6). The spatial association between the two species were generally positive with significant positive associations ($P < 0.025$) occurring frequently during the period but not continuously for more than two weeks (Table 6).

During the late season, the weekly spatial pattern of *C. quilicii* females mostly had an aggregated spatial pattern ($I_a > 1$) but showed no significant aggregation from week 6 to week 25 ($0.025 < P_a < 0.975$) (Table 7). *Ceratitis capitata* females showed a weekly aggregated spatial pattern ($I_a > 1$) throughout the late season, with significant aggregation ($P_a < 0.025$) occurring in some weeks. From week 12 to week 16, *C. capitata* females continuously showed a significant spatial pattern. During the late season, from week 10 to week 25, *C. capitata* and *C. quilicii* was significantly spatially associated ($P < 0.025$). Weekly spatial pattern-and association indices with their respective probability levels for the late season are contained in Table 7.

Table 6. The spatial patterns of *Ceratitis capitata* and *Ceratitis quilicii* females and the spatial association between the two species in each sampling week of the early season (week 35 to week 5) on Welgevallen, Stellenbosch, South Africa.

	Spatial pattern				**Spatial association**	
	Ceratitis capitata		*Ceratitis quilicii*		*Ceratitis capitata/Ceratitis quilicii*	
Week	I_a	P_a	I_a	P_a	X	P
Week 35	0.859	0.6793	-	-	-	-
Week 36	1.138	0.2096	1.141	0.1939	0.1965	0.1108
Week 37	1.094	0.2317	0.741	0.9591	0.1760	0.1646
Week 38	-	-	1.366	0.1269	-	-
Week 39	1.366	0.1269	1.290	0.0994	**0.6486**	**0.0029**
Week 40	1.366	0.1269	1.366	0.1269	1	
Week 41	-	-	1.490	0.0634	-	-
Week 42	-	-	1.151	0.1916	-	-
Week 43	-	-	**1.750**	**0.0085**	-	-
Week 44	-	-	1.231	0.1260	-	-
Week 45	1.223	0.1341	1.433	0.0436	**0.6654**	**0.0001**
Week 46	1.236	0.1021	1.342	0.0562	-0.0254	0.5833
Week 47	1.290	0.1080	1.091	0.2542	0.2598	0.0388
Week 48	0.979	0.4534	1.123	0.2006	**0.6528**	**0.0011**
Week 49	1.251	0.1291	**1.639**	**0.0103**	**0.6145**	**0.0001**
Week 50	1.143	0.2029	1.279	0.1237	0.0709	0.2757
Week 51	1.175	0.1583	**1.599**	**0.0144**	0.1351	0.1316
Week 52	**1.883**	**0.0076**	**1.747**	**0.0081**	0.0922	0.2885
Week 3	**1.700**	**0.0022**	**1.747**	**0.0081**	**0.6658**	**0.0001**
Week 4	**1.552**	**0.0198**	1.584	0.0319	**0.3749**	**0.0010**
Week 5	1.130	0.2006	1.121	0.2290	0.2295	0.0336

I_a = index of aggregation where $I_a = 1$, random spatial pattern; $I_a < 1$, regular spatial pattern and $I_a > 1$, aggregated spatial pattern.
P_a = probability level of aggregation index where $P_a > 0.975$, regular spatial arrangement (underlined); $P_a < 0.025$, aggregated spatial arrangement (bold) and $0.025 < P_a < 0.975$, random spatial arrangement.
X = index of association where $X > 0$, association and $X < 0$ dissociation
P = probability level of association index, where $P < 0.025$, significant association (**bold**)

Table 7. The spatial patterns of *Ceratitis capitata* and *Ceratitis quilicii* females and the spatial association between the two species in each sampling week of the late season (week 6 to week 25) on Welgevallen, Stellenbosch, South Africa.

	Spatial pattern				**Spatial association**	
	Ceratitis capitata		*Ceratitis quilicii*		*Ceratitis capitata/Ceratitis quilicii*	
Week	I_a	P_a	I_a	P_a	X	P
Week 6	1.632	0.0139	1.183	0.1610	0.3045	0.0270
Week 7	1.362	0.0481	0.902	0.6478	0.1438	0.1450
Week 8	1.209	0.1493	1.087	0.2596	0.1921	0.0716
Week 9	1.403	0.0414	1.213	0.1156	0.2796	0.0311
Week 10	**1.628**	**0.0054**	1.118	0.2276	**0.4760**	**0.0003**
Week 11	1.306	0.0742	1.246	0.1278	**0.5923**	**0.0001**
Week 12	**1.654**	**0.0076**	1.402	0.0657	**0.7285**	**0.0001**
Week 13	**1.526**	**0.0229**	1.266	0.1053	**0.7049**	**0.0001**
Week 14	**1.900**	**0.0004**	1.121	0.2173	**0.6567**	**0.0001**
Week 15	**1.500**	**0.0153**	1.173	0.1777	**0.7444**	**0.0001**
Week 16	**1.536**	**0.0135**	1.278	0.0900	**0.6984**	**0.0002**
Week 17	1.385	0.0495	1.256	0.1053	**0.7565**	**0.0001**
Week 18	1.309	0.0931	1.041	0.3401	**0.7209**	**0.0001**
Week 19	1.406	0.0432	1.323	0.0612	**0.7781**	**0.0001**
Week 20	**1.483**	**0.0211**	1.242	0.1170	**0.6496**	**0.0003**
Week 21	1.437	0.0297	1.293	0.0679	**0.5619**	**0.0010**
Week 22	1.316	0.0585	1.067	0.2834	**0.6319**	**0.0001**
Week 25	**1.954**	**0.004**	1.428	0.0337	**0.4997**	**0.0001**

I_a = index of aggregation where $I_a = 1$, random spatial pattern; $I_a < 1$, regular spatial pattern and $I_a > 1$, aggregated spatial pattern.
P_a = probability level of aggregation index where $P_a > 0.975$, regular spatial arrangement (underlined); $P_a < 0.025$, aggregated spatial arrangement (bold) and $0.025 < P_a < 0.975$, random spatial arrangement.
X = index of association where $X > 0$, association and $X < 0$ dissociation
P = probability level of association index, where $P < 0.025$, significant association (bold)

Discussion

The results showed that *C. capitata* and *C. quilicii* females were spatially associated, most prominently in home gardens, natural vegetation, citrus and nectarines. This suggests a niche overlap in terms of host utilisation between the two species. Later in the season consistently more significant associations between the two species were observed than early in the season, suggesting a strong influence of ripening fruit in the late season (see harvest dates in Figure 4). This is supported, by the damage assessment results where both species were reared from fruits obtained from the above-mentioned habitat types, however, no comment can be made regarding natural vegetation, as fruits were not collected from these areas. Both citrus and nectarines have been found to be hosts to *C. capitata* and *C. quilicii* (assumed from records of "*C. rosa"* from the Western Cape), while kei-apple and loquat found in home gardens have been also found to be good hosts for both species (DeVilliers *et al.* 2013).

Ceratitis capitata and *C. rosa* on La Réunion have been shown to segregate ecologically and geographically (Duyck *et al.* 2006b), however, *C. rosa* populations from La Réunion exclusively belongs to the 'R2' morphotype (Virgilio *et al.* 2013), which was later described as a new species, *C. quilicii* (De Meyer *et al.* 2016). Duyck *et al.* (2006), suggest that the niche segregation between *C. capitata* and *C. quilicii* on La Réunion is largely due to niche-dependent competition based on climate. However, this spatial segregation was not observed on the scale, which this study was conducted. Strong spatial associations were observed between the two species. Although the species were spatially associated, there was a temporal difference in their spatial patterns. Both species had aggregated spatial patterns, however, *C. capitata* significantly aggregated more towards the end of the season while *C. quilicii* significantly aggregated more towards the beginning of the season. This suggests the influence of temperature on the temporal patterns of these flies. Findings on

the lower developmental thresholds for *C. quilicii* (7.74 °C) (Tanga *et al.* 2015) suggests that *C. quilicii* is better adapted to lower temperatures compared to *C. capitata* (10.5 °C) (Grout & Stoltz 2007). *Ceratitis quilicii* would have started developing earlier than *C. capitata* and therefore resulted in early clustering of high trap catches followed by *C. capitata.* Temperature records show increasing temperatures from cold winter temperatures at which point *C. quilicii* would have started their reproductive cycle, to hotter summer temperatures during the time, that *C. capitata* peaked. Note unusual rainfall patterns, and drought in the Western Cape, during the time of the study (Botai *et al.* 2017), could have influenced the data. Temperature and rainfall parameters are investigated in detail in Chapter 3.

Furthermore, other factors such as location of host fruits and their phenology have been found to influence the spatial distribution of *C. capitata* (Katsoyannos *et al.* 1998; Papadopoulos *et al.* 2001a, 2003; Vera *et al.* 2002; Sciarretta & Trematerra 2011). Therefore, these factors might have contributed to the temporal differences in the spatial pattern and trap catches observed, as the orchard in which this study was conducted, consisted of mixed fruit with varying fruit ripening periods. Furthermore, the existence of larval and adult competition between the two species (Duyck *et al.* 2006a), also might have influenced the temporal differences between the two species.

It was clear that the natural vegetation, dominated by wild olives, *Eucalyptus* spp. and *Podocarpus* spp., was a suitable niche for *C. quilicii* females based on the trap catches, suggesting that they use natural vegetation as a refuge. While it is suggested, that wild olives may not be optimal hosts for *C. capitata* (Mkize 2008), very few studies have been done on host status of olives in South Africa for *Ceratitis* spp. Therefore, further research to determine natural hosts of *C. quilicii* in the Western Cape are important to understand the ecology and dynamics of this fly in order to contribute to management it as part of an

integrated pest management strategy. Furthermore, the microclimate within the natural vegetation, comprised of big *Eucalyptus* spp., creating shade, as well as possible alternate food sources in the natural vegetation, which might have favoured *C. quilicii* females (see also Israely *et al.* 1997).

The spatial patterns of the two species were similar in that they mostly formed patches in home gardens, natural vegetation and on the edge of the orchard bordering home gardens, while gaps mostly occurred within the orchard. According to Aluja (1996), tephritid fruit flies typically invade commercial fruit orchards from natural areas or unmanaged fruit trees in close vicinity outside of these. Israely *et al.* (1997) also reported this movement of *C. capitata* from home gardens to commercial orchards in central Israel. Myburgh (1956) highlighted the important role, which unmanaged home garden hosts play in harbouring fruit flies during times when no ripe fruits are available in commercial orchards, until fruits in commercial orchards become susceptible to fruit fly attack. The patches formed in home gardens and on the edge of the orchard support the findings of the previous authors, i.e. of movement from home gardens and natural vegetation to commercial orchards, indicated by the patches on the edge of the orchard.

Cohen and Yuval (2000) found that trapping on the perimeter (± 30 traps per 350 trees) of commercial plum, pear and persimmons orchards in Israel, bordering unmanaged fruit trees, resulted in acceptable control of *C. capitata* in these orchards. Furthermore, the fruit orchards in our study area was well managed and during the time of the study was under intensive fruit fly control, which included weekly bait sprays starting in August of 2016, while fruit flies in home gardens and natural vegetation were not actively managed throughout the season. These factors might have contributed to the spatial patterns observed. Furthermore, the trapping density in the orchard, which was denser compared to the habitats outside of

the orchard, might have influenced the spatial patterns observed. However, further research needs to be conducted on the impact of trapping density on the spatial patterns of fruit flies.

Based on the results from this study it is recommended that unmanaged fruit in home gardens should be stripped to prevent population build-up in these habitats, which could later invade high value export fruit crops in commercial orchards. Based on the findings of Cohen and Yuval (2000), perimeter trapping could aid in controlling fruit flies in commercial orchards bordering unmanaged host fruits. According to Papadopoulos *et al.* (2001), early season fruit should be removed (which may not viable at the Welgevallen orchard), or be more heavily controlled, to prevent early season population build-up. Leza *et al.* (2008), found that female mass trapping (50 traps/ha), supplemented with bait application are more effective in reducing *C. capitata* females in citrus orchards in Mallorca, Spain, than when only applying bait sprays, when compared to the same insecticide. It is assumed that this method would work for *C. capitata* and *C. quilicii* in other fruit kinds as well, but would need to be tested. Furthermore, releases of sterile *C. capitata* males in early maturing hosts could also aid the management of *C. capitata* and prevent the populations reaching high population peaks later in the season, as SIT is most effective when the adult populations are low and a high sterile/wild male ratio is achieved (Dyck *et al.* 2005; Barnes 2009). In addition, James *et al*. (2018) showed that the local entomopathogenic nematode (EPN), *Heterorhabditis noenieputensis* Malan, Knoetze and Tiedt (Rhabditida: Heterorhabditidae), could be used to control third instar *C. capitata* larvae entering the ground before pupation. This could be ideal for late season application in orchards to target the pupating larvae, before winter, possibly reducing the following spring populations. However, this EPN is not currently commercially available and further field trials would need to be conducted to establish the efficacy of this control method, particularly regarding timing of the application.

Further research should focus on *C. quilicii,* as no information exists for biological control using entomopathogens against this species.

With the introduction of *B. dorsalis* in the Western Cape, the abundance and distribution of *C. capitata* might be influenced due to competitive displacement as shown by Ekesi *et al.* (2009) between *C. cosyra* and *B. dorsalis* in Kenya. This could be true for *C. capitata* and *B. dorsalis* in the Western Cape, where *B. dorsalis* might displace *C. capitata* to specific niches, due to its strong competitive traits, highlighted by Duyck *et al.* (2004). *Bactrocera dorsalis* displaced *C. capitata* from low elevation areas to high elevation areas in Hawaii, but was still able to utilise coffee in low elevation areas from which it was displaced (Duyck *et al.* 2004). Displacement of *C. capitata* by *B. dorsalis* was suggested by Mwatawala *et al.* (2015), and *B. dorsalis* has dominated *Ceratitis* species in Tanzania after its introduction, in terms of abundance, host range and infestation rate (Mwatawala *et al.* 2009). Therefore, if *B. dorsalis* becomes established in the Western Cape, it would offer an opportunity to study the competitive displacement properties of *B. dorsalis* on *C. capitata* as the distribution and abundance of *C. capitata* in the Western Cape is well documented. Currently there is no literature citing *B. dorsalis* displacing *C. quilicii*.

The other species caught during the study period do not pose a current threat, due to their low abundance and the fact that they have not been sampled from the fruits collected.

Conclusion

This study indicated that *C. capitata* and *C. quilicii* females spatially share resources in a heterogeneous environment on an orchard scale, but that their spatial temporal patterns differ to some degree, with *C. quilicii* significantly aggregating early in the season, while *C. capitata* significantly aggregates towards the end of the season. These spatial patterns are

determined by biotic and abiotic factors, including the physiological responses of these flies to temperature and rainfall, interspecific competition, host fruit location and fruiting phenology. Unmanaged home gardens sustained both species throughout the season while *C. quilicii* also found a suitable refuge in natural vegetation. The results also suggest dispersal of the two species from unmanaged home gardens and natural vegetation to commercial orchards, based on the aggregation of these flies in these habitats as well as on the edge of the commercial fruit orchard, facing these habitats. Although, this is the first study investigating the spatial patterns and associations between *C. capitata* and *C. quilicii,* further experiments would need to be conducted to show if the spatial patterns and associations observed are consistent across seasons and years.

References

ALUJA, M. 1996. Future trends in fruit fly management. In: Aluja, M. & Liedo, P. (Eds.) *Fruit Flies: Biology and Management*. 309–320. Springer, New York.

BARNES, B.N. 2009. The Sterile Insect Technique (S.I.T.) for management of Mediterranean Fruit Flies in fruit production - A technical manual. Stellenbosch, South Africa.

BARR, N.B. & MCPHERON, B.A. 2006. Molecular phylogenetics of the genus *Ceratitis* (Diptera: Tephritidae). *Molecular Phylogenetics and Evolution* **38**(1): 216–230.

BATEMAN, M.A. 1972. The ecology of fruit flies. *Annual Review of Entomology* **17**: 493–518.

BOTAI, C.M., BOTAI, J.O., DE WIT, J.P., NCONGWANE, K.P. & ADEOLA, A.M. 2017. Drought Characteristics over the Western Cape Province, South Africa. *Water* **9**(879): 1–16.

BROWN, L. & PRINGLE, K. 2006. Monitorstelsel vir plae van kernvrugte. Stellenbosch.

COHEN, H. & YUVAL, B. 2000. Perimeter trapping strategy to reduce Mediterranean fruit fly (Diptera: Tephritidae) damage on different host species in Israel. *Journal of economic entomology* **93**(1): 721–725.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DUTILLEUL, P., CLIFFORD, P., RICHARDSON, S. & HEMON, D. 1993. Modifying the t Test for Assessing the Correlation Between Two Spatial Processes. *Biometrics* **49**(1): 305–314.

DUYCK, P.F., DAVID, P., JUNOD, G., BRUNEL, C., DUPONT, R. & QUILICI, S. 2006a. Importance of competition mechanisms in successive invasions by polyphagus tephritids in La Réunion. *Ecology* **87**(7): 1770–1780.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2004. A review of relationships between interspecific competition and invasions in fruit flies (Diptera: Tephritidae). *Ecological Entomology* **29**(5): 511–520.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006b. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

DYCK, V.A., HENDRICHS, J. & ROBINSON, A.S. 2005. *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. *Springer, Dordrecht, The Netherlands*.

EKESI, S., BILLAH, M.K., NDERITU, P.W., LUX, S.A. & RWOMUSHANA, I. 2009. Evidence for Competitive Displacement of *Ceratitis cosyra* by the Invasive Fruit Fly *Bactrocera invadens* (Diptera: Tephritidae) on Mango and Mechanisms Contributing to the Displacement. *Journal of Economic Entomology* **102**(3): 981–991.

GROUT, T.G. & STOLTZ, K.C. 2007. Developmental Rates at Constant Temperatures of Three Economically Important *Ceratitis* spp. (Diptera: Tephritidae) From Southern Africa. *Environmental Entomology* **36**(6): 1310–1317.

HENDRICHS, J., TERESA VERA, M., DE MEYE, M. & CLARKE, A.R. 2015. Resolving cryptic species complexes of major tephritid pests. *ZooKeys* **2015**(540): 5–39.

ISRAELY, N., YUVAL, B., KITRON, U. & NESTEL, D. 1997. Population fluctuations of adult Mediterranean fruit flies (Diptera: Tephritidae) in a Mediterranean heterogeneous agricultural region. *Environmental Entomology* **26**(6): 1263–1269.

ISRAELY, N., ZIV, Y. & GALUN, R. 2005a. Metapopulation Spatial – Temporal Distribution Patterns of Mediterranean Fruit Fly (Diptera: Tephritidae) in a Patchy Environment. *Annals of the Entomological Society of America* **98**(3): 302–308.

ISRAELY, N., ZIV, Y. & OMAN, S.D. 2005b. Spatiotemporal Distribution Patterns of Mediterranean Fruit Fly (Diptera : Tephritidae) in the Central Region of Israel. *Annals of the Entomological Society of America* **98**(1): 77–84.

JAMES, M., MALAN, A.P. & ADDISON, P. 2018. Surveying and screening South African entomopathogenic nematodes for the control of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann). *Crop Protection* **105**: 41–48. Elsevier.

KARSTEN, M., BARTON, M., ADDISON, P., ADDISON, M.F. & TERBLANCHE, J.S. 2018. A synthesis for managing invasions and pest risks simultaneously for tephritid fruit flies in South Africa. *Entomologia Experimentalis et Applicata*(2): 1–13.

KATSOYANNOS, B.I., KOULOUSSIS, N.A. & CAREY, J.R. 1998. Seasonal and annual occurrence of Mediterranean fruit flies (Diptera: Tephritidae) on Chios Island, Greece: differences between two neighboring citrus orchards. *Annals of the Entomological Society of America* **91**(1): 43–51.

LEZA, M.M., JUAN, A., CAPLLONCH, M. & ALEMANY, A. 2008. Female-biased mass trapping vs . bait application techniques against the Mediterranean fruit fly, *Ceratitis capitata* (Dipt., Tephritidae). *Journal of Applied Entomology* **132**(9–10): 753–761.

MANRAKHAN, A. & ADDISON, P. 2014. Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science* **70**(4): 651–660.

DE MEYER, M., DELATTE, H., EKESI, S., JORDAENS, K., KALINOVÁ, B., MANRAKHAN, A., MWATAWALA, M., STECK, G., VAN CANN, J., VANÍČKOVÁ, L., BŘÍZOVÁ, R. & VIRGILIO, M. 2015. An integrative approach to unravel the *Ceratitis* FAR (Diptera, tephritidae) cryptic species complex: A review. *ZooKeys* **2015**(540): 405–427.

DE MEYER, M., MWATAWALA, M., COPELAND, R.S. & VIRGILIO, M. 2016. Description of new *Ceratitis* species (Diptera: Tephritidae) from Africa, or how morphological and DNA data are complementary in discovering unknown species and matching sexes. *European Journal of Taxonomy* **233**: 1–23.

MKIZE, N. 2008. Insect pests of cultivated and wild olives, and some of their natural enemies, in the Eastern Cape, South Africa. Rhodes University.

MWATAWALA, M.W., DE MEYER, M., MAKUNDI, R.H. & MAERERE, A.P. 2009. An overview of *Bactrocera* (Diptera: Tephritidae) Invasions and Their Speculated Dominancy over Naive Fruit Fly Species in Tanzania. *Journal of Entomology* **6**(1): 18–27.

MWATAWALA, M.W., VIRGILIO, M., JOSEPH, J. & DE MEYER, M. 2015. Niche partitioning among two *Ceratitis rosa* morphotypes and other *Ceratitis* pest species (Diptera, Tephritidae) along an altitudinal transect in Central Tanzania. *ZooKeys* **540**: 429–442.

MYBURGH, A.C. 1956. Bionomics and control of the fruit flies, *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.), in the Western Cape. University of Stellenbosch.

NESTEL, D., KATSOYANNOS, B.I., NEMNY-LAVY, E., MENDEL, Z. & PAPADOPOULOS, N.T. 2002. Spatial analysis of Medfly populations in heterogeneous landscapes. *Proceedings of 6th International Fruit Fly Symposium*. 35–43. Stellenbosch, South Africa.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I., CAREY, J.R. & KOULOUSSIS, N.A. 2001. Seasonal and annual occurrence of the Mediterranean fruit fly (Diptera: Tephritidae) in northern Greece. *Annals of the Entomological Society of America* **94**(1): 41–50.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PERRY, J., WINDER, L., HOLLAND, J. & ALSTON, R. 1999. Red-blue plots for detecting cluster in count data. *Ecology Letters* **2**(May 2016): 106–113.

PERRY, J.N. 1995. Spatial Analysis by Distance Indices. *The Journal of Animal Ecology* **64**(3): 303.

PERRY, J.N. 1998. Measures of spatial pattern for counts. *Ecology* **79**(3): 1008–1017.

PERRY, J.N., BELL, E.D., SMITH, R.H. & WOIWOD, I.P. 1996. SADIE: software to measure and model spatial pattern. *Aspects of Applied Biology* **46**: 95–102.

PERRY, J.N. & DIXON, P.M. 2002. A new method to measure spatial association for ecological count data. *Ecoscience* **9**(2): 133–141.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

TANGA, C.M., MANRAKHAN, A., DANEEL, J.H., MOHAMED, S.A., FATHIYA, K. & EKESI, S. 2015. Comparative analysis of development and survival of two Natal fruit fly *Ceratitis rosa* Karsch (Diptera, tephritidae) populations from Kenya and South Africa. *ZooKeys* **540**: 467–487.

TIBCO SOFTWARE INC. 2017. Statistica (data analysis software system).

VERA, M.T., RODRIGUEZ, R., SEGURA, D.F., CLADERA, J.L. & SUTHERST, R.W. 2002. Potential geographical distribution of the Mediterranean fruit fly, Ceratitis capitata (Diptera: Tephritidae), with emphasis on Argentina and Australia. *Environmental Entomology* **31**(6): 1009–1022.

VIRGILIO, M., BACKELJAU, T., BARR, N. & DE MEYER, M. 2008. Molecular evaluation of nominal species in the *Ceratitis fasciventris*, *C. anonae*, *C. rosa* complex (Diptera: Tephritidae). *Molecular Phylogenetics and Evolution* **48**(1): 270–280.

VIRGILIO, M., DELATTE, H., QUILICI, S., BACKELJAU, T. & DE MEYER, M. 2013. Cryptic diversity and gene flow among three African agricultural pests: *Ceratitis rosa*, *Ceratitis fasciventris* and *Ceratitis anonae* (Diptera, Tephritidae). *Molecular Ecology* **22**(9): 2526–2539.

WHITE, I.M. & ELSON-HARRIS, M.M. 1992. *Fruit flies of economic significance: their identification and bionomics*. CAB International/Australian Centre for International Agricultural Research, Wallingford, UK/Canberra.

WICHMANN, B.A. & HILL, I.D. 1982. Algorithm AS 183: An Efficient and Portable Pseudo-Random Number Generator. *Journal of the Royal Statistical Society* **31**(2): 188–190.

Chapter 3

The area-wide spatio-temporal distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) in a heterogeneous fruit production region of the Western Cape, South Africa

This chapter has been accepted with minor revisions in *African Entomology* as Bekker, G.F.H.v.G., Baard, N., Addison, M.F., Van Niekerk, A. and Addison, P. The area-wide spatio-temporal distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) in a heterogeneous fruit production region of the Western Cape, South Africa

Abstract

The area-wide spatio-temporal distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) trap catches was investigated in the sub-regions of Elgin/Grabouw, Villiersdorp and Vyeboom, collectively known as EGVV, Western Cape, South Africa. The primary fruit crops planted in EGVV, include various varieties and cultivars of pome fruit, stone fruit, wine grapes and berries. Fruit fly trapping data from four consecutive seasons were sourced from a centralised database managed by a non-profit company, which operates an area-wide integrated pest management (AW-IPM) programme, incorporating the sterile insect technique (SIT) for *C. capitata* fruits flies, within EGVV. Mean trap catch between regions and seasons was compared. The spatio-temporal variation of mean trap catch was also investigated. A geographic information system (GIS) was used to symbolise and map mean trap catch per trap location. There were no significant differences in the mean trap catch between regions. Significant differences did occur between some seasons in Villiersdorp and Vyeboom but not in Elgin/Grabouw. The spatial distribution of trap catches between seasons varied greatly, indicating no definitive pattern. However, traps with the highest seasonal trap catch occurred mostly in the south-eastern parts of each region and traps with lower levels were distributed mostly in the north-western parts of each region. The results suggest a relationship between the geographic characteristics of EGVV and the abundance and distribution of *C. capitata*. It is recommended that an in-depth spatio-temporal analysis be carried out to better understand the geographical distribution of trap catches in the study area.

Introduction

A successful, area-wide integrated pest management (AW-IPM) programme requires a clear understanding of the spatio-temporal distribution of the target pest in relation to its environment (Cox & Vreysen 2005). Geographic information systems (GIS), together with spatial analysis methods, allow researchers to study and better understand the large-scale spatial structure and dynamics of insect populations and how they are influenced by heterogeneous environments (Dminic *et al.* 2010).

A number of studies, which spatially analysed insect distribution in fruit orchards and vineyards, including Lepidoptera and Diptera species, have been conducted (Sciarretta & Trematerra 2014). An important family within the Diptera order is tephritid fruit flies, of which some are key pests of commercial agricultural fruit and vegetable crops globally (White & Elson-Harris 1992). Tephritid fruit fly spatial distribution patterns, especially those of *Ceratitis capitata* (Wiedemann) and *Bactrocera oleae* (Rossi) have been the subject of numerous spatial analysis studies (Papadopoulos *et al.* 2003; Alemany *et al.* 2006; Kounatidis *et al.* 2008; Epsky *et al.* 2010; Sciarretta & Trematerra 2011). However, most of this research was done on a very small field scale. A notable exception is that of Sciarretta & Trematerra (2011), which was conducted in a ± 500 ha agricultural landscape near Rome in Italy and included mixed fruit orchards. However, this area is still relatively small compared to the areas on which most AW-IPM programmes are applied.

The aim of AW-IPM is to manage all individuals of the pest population in time and space (Hendrichs *et al.* 2015a). The areas over which AW-IPM actions are applied generally cover large geographic regions (Lindquist 2000). However, there has been a shift towards a new paradigm of area-wide pest management that intends to manage the target pest populations more precisely by following their spatio-temporal development and their damage (Nestel *et*

al. 2004). For instance, Enkerlin *et al.* (2016) used trapping data obtained from an AW-IPM programme spanning more than 80 000 km^2 in Guatemala to investigate the effect of biotic and abiotic factors on the population dynamics of *C. capitata*. The aim of their study was to understand the interactions between the pest population and the environmental factors, although the spatio-temporal distribution of the flies has not been described. Midgarden and Lira (2006) used trapping data from the same programme to investigate the ecological relationship between *C. capitata* populations and the phenology of coffee plants, again to improve the understanding of the dynamics of this pest in relation to its environment. Furthermore, Kounatidis *et al.* (2008), investigated the effect of elevation on the spatio-temporal distribution of *B. oleae*, using data sourced from an AW-IPM programme in Northern Greece. These studies showed that pest population data, from AW-IPM programmes, are valuable for generating information regarding long-term spatio-temporal relationships between the pest and its environment, which can be used for managing pests more precisely in large geographic regions.

The employment of GIS and spatial analyses in AW-IPM studies have assisted in elucidating the spatio-temporal distribution of the target pest in relation to its environment and assist managers to make more informed management decisions at different locations and at different spatial scales (Kitron *et al.* 1996; Beckler *et al.* 2005; Cox & Vreysen 2005; Guidotti *et al.* 2005; Castrignanò *et al.* 2012; Midgarden *et al.* 2014). For example, Castrignanò *et al.* (2012) used GIS to delineate monitoring zones to optimise fruit fly monitoring efforts, while Guidotti *et al.* (2005) created a regional map of where *B. oleae* infestation starts in Tuscany (Italy), which growers could use to time-control actions in specific locations. Furthermore, in Mexico, Guatemala, Belize and the United States, a coordinated AW-IPM programme against *C. capitata* employed GIS/GPS technology, to identify population hot spots, and focus their management activities (McGovern *et al.* 2008). These studies are not

limited to fruit flies. Kitron *et al.* (1996) spatially analysed tsetse fly distributions in the Lambwe Valley in Kenya, using large-scale trapping data together with satellite imagery and GIS, to identify factors associated with local variations of fly density and to describe their spatial distribution. The outcome of such analyses can, therefore, contribute to the effective planning and implementation of AW-IPM actions (Papadopoulos *et al.* 2003; Hendrichs *et al.* 2007a; Enkerlin *et al.* 2016).

The relationship between the spatio-temporal distribution of insect pests and their environment is complex, with interactions of multiple environmental factors and intrinsic interactions of the species playing a role in determining how insects vary over time and space (Berryman 2003; Papadopoulos *et al.* 2003). Trapping networks that generate spatio-temporal pest data form an integral part of any AW-IPM programme (IAEA 2003). These information-rich datasets present the opportunity for researchers to analyse and describe pest population trends across multiple seasons, and over large geographic regions consisting of various habitat and crop types, climatic conditions and topographies.

In South Africa, *C. capitata* is one of the major, polyphagous, fruit fly pests of commercially produced fruit, especially the Western Cape (Mumford & Tween 1997; Barnes *et al.* 2007). In 1997, a pilot AW-IPM project against *C. capitata* was launched in the Hex River Valley, a geographically isolated table grape production region in the Western Cape. This pilot project only focussed on fruit fly monitoring and the Sterile Insect Technique (SIT). Since 1997, the project has developed into a commercial AW-IPM programme, incorporating SIT and the Bait Application Technique (BAT) (Barnes *et al.* 2015). Today the programme operates in three main fruit production regions across the Western Cape, including the Hex River Valley (-33.469689°, 19.645225°); Warm Bokkeveld (-33.364378°, 19,311459°) (including Tulbagh and Wolseley); and the regions of Elgin/Grabouw (-34.154536°, 19.029693°) Villiersdorp (-

33.991976°, 19.289763°) and Vyeboom (-34.067610°, 19.109035), collectively known as EGVV (Figure 1) (Barnes *et al.* 2015). The Villiersdorp sub-region has been excluded from this AW-IPM programme since July 2016.

Israely *et al.* (2005a) concluded that it is very difficult to explain *C. capitata* population distributions within the landscape without fully understanding their large-scale spatio-temporal dynamics. Despite the spatial distribution studies conducted on an area-wide scale on fruit flies, little is known about how *C. capitata* populations vary over time and space in large heterogeneous agricultural systems. This chapter is an initial exploration of the available area-wide *C. capitata* trapping data, andinvestigating how *C. capitata* populations vary over space and time, with the aim of mapping, quantifying and describing the spatio-temporal distributions of trap catches. This was conducted by visualising fruit fly trap catches obtained from an AW-IPM programme and symbolising the trap catches using subjective management thresholds. A GIS and spatial analysis were used to produce a series of maps, which are interpreted within the context of making more targeted management recommendations to improve AW-IPM of this pest.

Material and methods

Study area

This study was conducted in EGVV, which is a heterogeneous agricultural production area, with fruit being the primary agricultural commodity (± 148 sq/km of commercial plantings). The primary fruit crops include various varieties and cultivars of pome fruit, stone fruit, citrus, wine grapes and berries. Pome fruit is the most widely planted crop, while citrus is grown in a relatively small area (Table 1; Figure 1). Even though pome fruit is the predominant crop planted in all of the regions, there are some differences in crop composition between the regions. Notable differences include stone fruit plantings in the Villiersdorp region, which

cover about 58% of the total area of stone fruit planted in EGVV (Table 1), with commercial fruit orchards also occurring near the town of Villiersdorp (Figure 1).

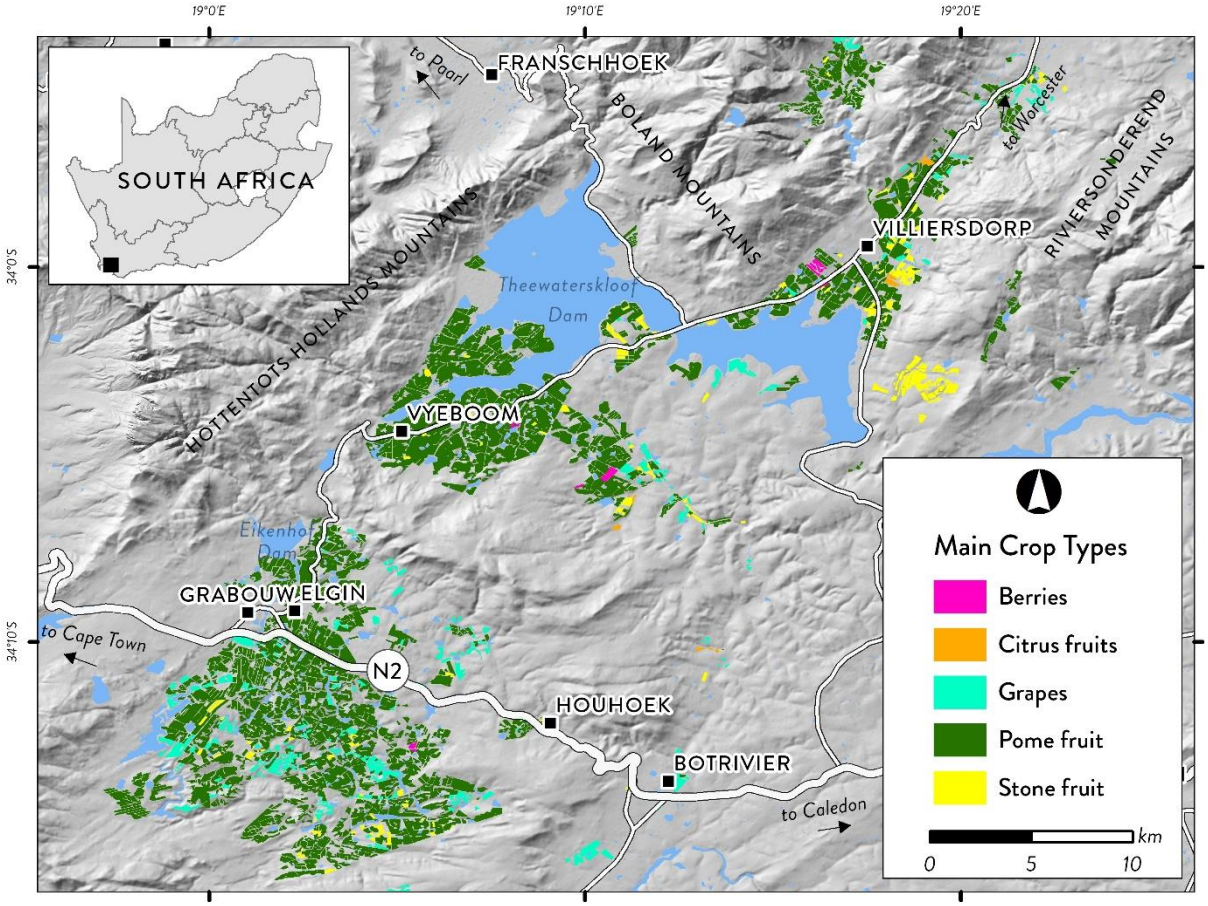

Figure 1. The spatial distribution of the main crop types planted in Elgin/Grabouw, Villiersdorp and Vyeboom, based on the Western Cape Department of Agriculture's 2013 crops census (Western Cape Department of Agriculture 2014).

The stone fruit planted in Villiersdorp and Vyeboom is dominated by peach and plum cultivars, interspersed with apricots (Western Cape Department of Agriculture 2014). Another important difference is grape (mainly wine grapes) plantings in Elgin/Grabouw, which covers 72% of the total area planted under grapes in EGVV. Elgin/Grabouw has relatively few stone fruit orchards compared to Villiersdorp (Table 1 and Figure 1) (Western Cape Department of Agriculture 2014).

Table 1. The total area (ha) of different fruit crops planted in Elgin/Grabouw, Villiersdorp and Vyeboom, based on the Western Cape Department of Agriculture's 2013 Flyover spatial dataset (Western Cape Department of Agriculture 2014).

Crop type	Elgin/Grabouw	Vyeboom	Villiersdorp	Total (ha)
Berries	10	41	33	84
Citrus	--	10	34	45
Grapes	854	192	142	1188
Pome fruit	6111	4005	2073	12189
Stone fruit	157	178	473	807

Fruiting phenology in EGVV varies according to crop type, variety and cultivar. The fruiting period stretches over all months of the year, with winter crops, like citrus and berries, bearing fruit from April to July and August to February respectively (Figure 2). Alternate fruit fly hosts in home- and farm gardens, such as in *Citrus* spp., loquat *Eriobotrya japonica* (Thunb.) Lindley, kei-apple *Dovyalis caffra* (J.D.Hook & Harvey) J. D.Hook, and guava *Psidium guajava* L., all bear fruit between March and October and thus fall outside the commercial harvest window (Manrakhan & Addison 2014). Furthermore, alternate hosts in natural areas, such as the invasive bugweed *Solanum mauritianum (*Scop.) and brambles *Rubus spp.*, also play an important role in sustaining fruit fly populations outside of orchards (Myburgh 1956; Barnes 2009). Consequently, although the main fruiting period for deciduous fruit is from November to May, potential hosts for fruit fly populations are present throughout the year. Home gardens and agricultural lands are therefore primary targets for fruit fly management programmes (Manrakhan & Addison 2014; Barnes *et al.* 2015). Indigenous vegetation, however, in this region comprises fynbos, characterised by Proteaceae, Ericaceae and Restionaceae, which is the highly endemic natural habitat of the Cape Floral Kingdom (CFR) and is a biodiversity hotspot with high conservation value (Myers *et al.* 2000). Karsten (2014) suggested that the relatively low number of fruit flies in the natural vegetation of the study area might be due to a lack of fleshy fruits native to fynbos.

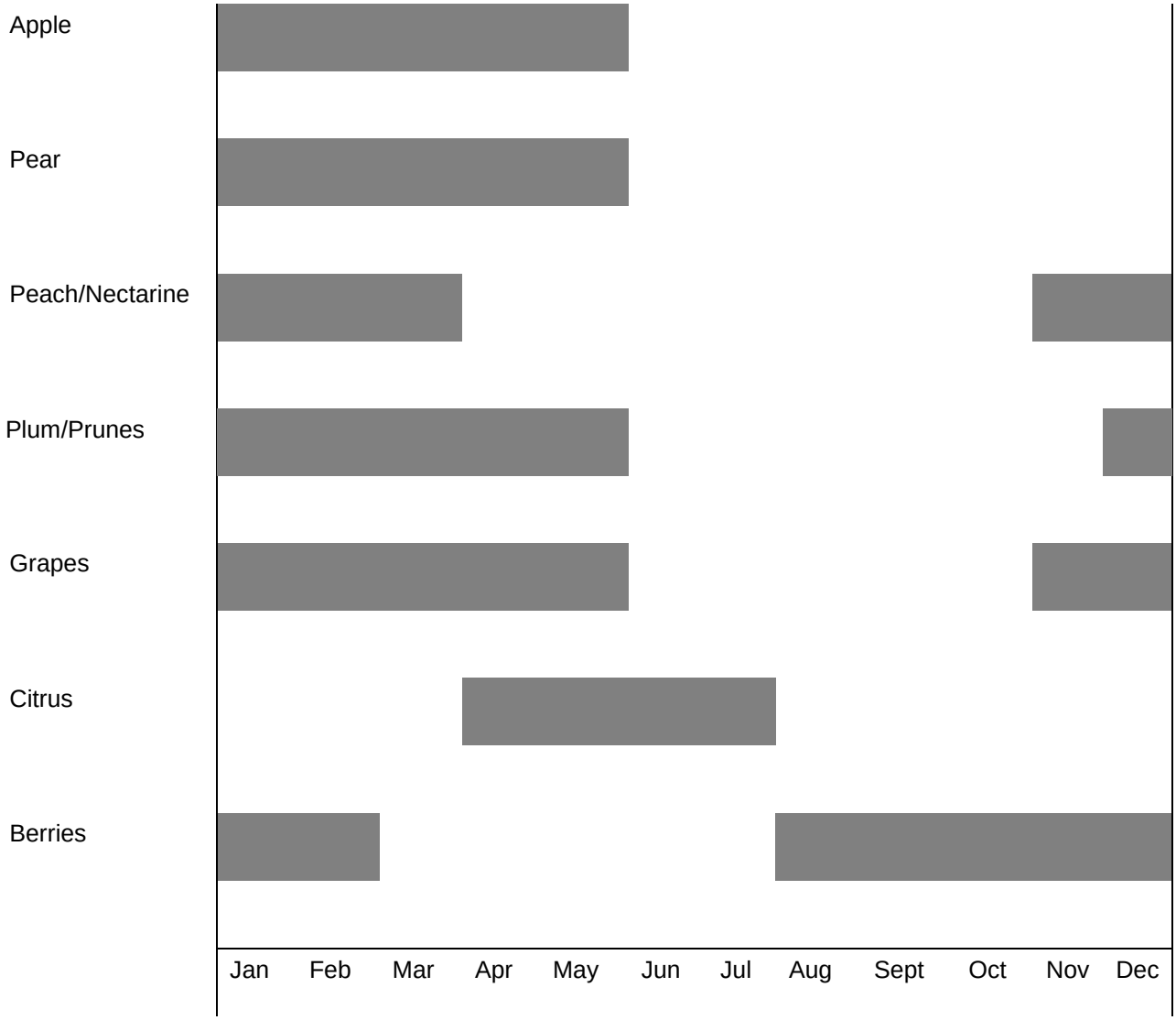

Figure 2. The fruiting phenology of the main crop types grown in Elgin/Grabouw, Villiersdorp and Vyeboom (adapted from Manrakhan & Addison 2014).

Climate

EGVV has a Mediterranean climate with hot, dry summers and cold, wet winters, with variable temperature and rainfall between the three regions (Figure 3). Elgin/Grabouw has a long-term mean annual rainfall of 944 mm and a long-term mean annual temperature of 14.9°C (Schulze 2006). The hottest month in Elgin/Grabouw is February with a long-term mean maximum temperature of 25.8°C, while the coldest month is July with a long-term

mean minimum temperature of 5°C. Vyeboom is drier, with a long-term mean annual rainfall of 768 mm and mean annual temperature of 15.4°C (Schulze 2006). The hottest month in Vyeboom is February with a long-term mean maximum temperature of 27.1°C, while the coldest month is July with a long-term mean minimum temperature of 5.3°C. Villiersdorp is the driest and hottest of the three regions, with a long-term mean annual rainfall and temperature of 568 mm and 16.5°C, respectively (Schulze 2006). The hottest month in Villiersdorp is February with a long-term mean maximum temperature of 28.3°C, while the coldest month is July with a long-term mean minimum temperature of 7.1°C.

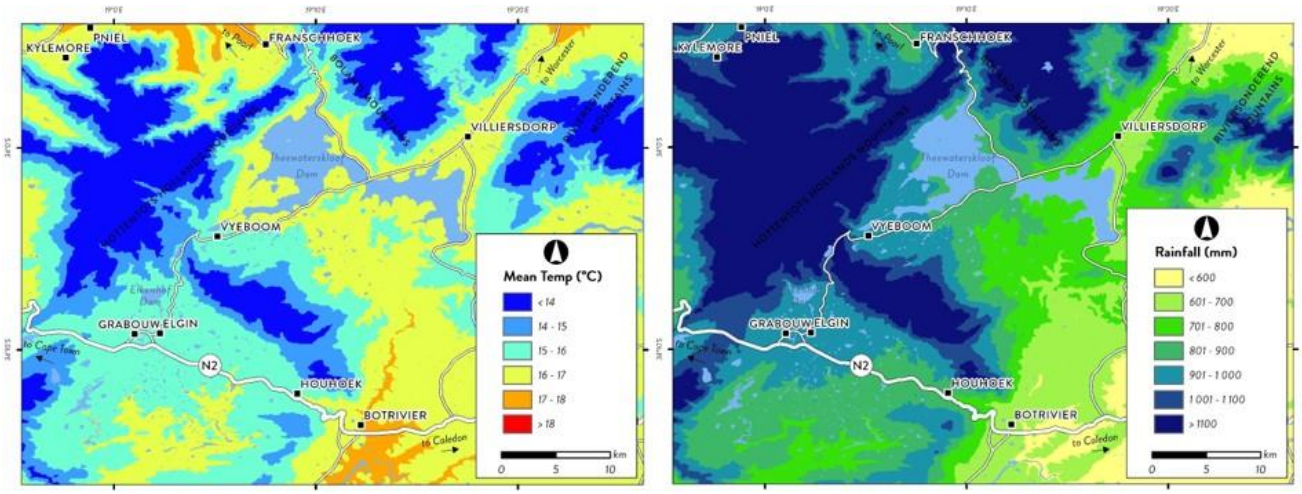

Figure 3. Long-term mean annual temperature (°C) (left) and rainfall (mm) (right) in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa.

Fruit fly management and monitoring

Before January 2004, the predominant control method for fruit flies in EGVV consisted of individual farmers applying routine ground bait sprays, consisting of a mixture of protein hydrolysate (attractant) and pesticide (organophosphates). Bait sprays were applied in commercial fruit orchards on a weekly basis during summer and at two-to-three-week intervals in winter months (Barnes & Venter 2006). An AW-IPM programme, incorporating SIT, was initiated in 2004 to manage *C. capitata* populations. The programme is operated by a private non-profit company, trading as FruitFly Africa (Pty) Ltd. (FFA). The management

approaches implemented by the programme initially included the bait application technique (BAT) (Roessler 1989), ground releases of *C. capitata* sterile males and cultural control (sanitation). These management efforts were mainly focused on areas with persistently high fruit fly populations, located outside of commercial plantings, including farm and urban gardens (Barnes & Venter 2006). During the study period, 9 000 – 38 000 sterile male flies were released per week per hectare throughout the year in these areas (Barnes & Venter 2006). The ground releases were conducted to target winter "breeding sites" to prevent large, early-season populations, which could potentially reinvade the commercial orchards when fruits start to ripen by the end of October and beginning of November (Barnes & Venter 2006).

Aerial BAT was included in the programme in 2010, with two applications applied between January and March at the beginning of the fruit harvest (pers. comm. Nando Baard). From the 2011/2012 season, four aerial BAT applications were applied between January and March. The main aim of the aerial BAT was to reduce fruit fly populations at the beginning of the harvest rather than trying to lower the high population levels that occur later in the season. In this way, the potentially high population pressure that could be experienced during mid-to-late season is reduced. A full account of the history of the FFA AW-IPM programme in the Western Cape, including EGVV, is detailed in Barnes *et al.* (2015).

FruitFly Africa is further responsible for the routine monitoring of fruit fly traps to evaluate fruit fly population levels in commercial plantings and in urban and farm home gardens. Trapping data are obtained using Chempac© Yellow Bucket traps (McPhail-type trap) with Vapona strips (dichlorvos) as retention system baited with three-component Biolure® Fruit Fly (Chempac Pty Ltd, Paarl, South Africa), which is a synthetic food-based lure (Ammonium acetate 211 g/kg; Trimethylamine hydrochloride 91 g/kg; 1,4-diaminobutane (Putrescine) 3

g/kg). Three-component biolure attracts both male and female *Ceratitis* spp., but is biased toward females (Heath *et al.* 1997; Epsky *et al.* 1999; Miranda *et al.* 2001; Leza *et al.* 2008). This trapping system is recommended for *C. capitata* AW-IPM programmes incorporating SIT to attract more wild females and fewer sterile and wild males (IAEA 2003; Ekesi *et al.* 2005).

The programme objectives dictate the placement of traps, with a recommended trapping density of one trap per 25ha. Trapping is conducted on a summer/winter cycle, with orchard traps activated (baited and monitored) during the fruiting and summer season (November to May), but deactivated (not baited) during the winter and early spring (June to October). When the data used in this study were collected, the summer season extended from November to May while winter and early spring extended from June to October. Subsequently, the programme managers changed these periods in order for the summer season to extend from October to June, while winter and early spring extend from July to September. In this paper, ‘season’ refers to the fruiting and summer season (November to May). Only urban and farm home garden traps are active during winter. All trap locations were georeferenced using a global positioning system (GPS) (Figure 4). Weekly trap counts and each trap location were recorded in a centralised Microsoft Excel dataset.

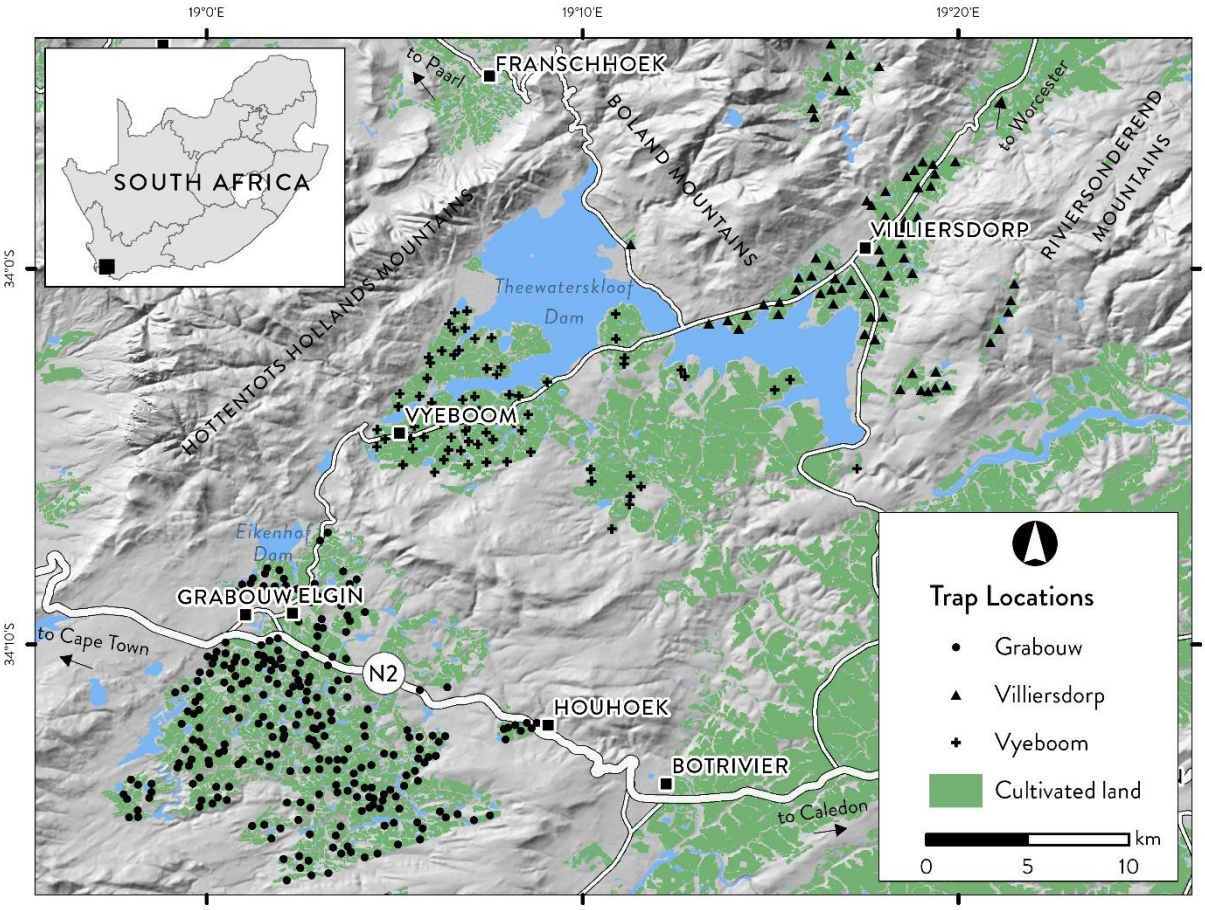

Figure 4. *Ceratitis capitata* trap locations in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa

Data analysis

Fruit fly trapping data and other related data for this study were obtained from FFA's centralised database. Viable traps were selected from the database using various selection criteria. The goal of the trap selection was to include as many viable traps as possible over consecutive seasons. A trap was considered viable if it complied with the following criteria:

- the location of the trap did not change for at least four consecutive seasons, and
- the trap was consistently monitored on a weekly to bi-weekly interval for the duration of the summer fruiting season (at least 30 weeks).

Available trapping data from the 2009/2010, 2010/2011, 2011/2012, 2012/2013, 2013/2014, 2014/2015 and 2015/2016 seasons were considered for the analysis. The selection criteria

contributed to standardising the data spatially and temporally, which assisted in the interpretation of the results.

Ceratitis capitata monitoring data from 399 traps, over four consecutive seasons (2010/2011, 2011/2012, 2012/2013 and 2013/2014), were selected for analyses. Of the selected traps, 86% were located in pome fruit orchards while the rest were distributed in stone fruit and citrus orchards, vineyards, home gardens and a berry plantation (Table 2). While these traps are associated with various fruit kinds, it is understood that no deductions of host utilisation can be made from trap cathes (Vargas *et al.* 1983a; Wong *et al.* 1983). Given the absence of fruit damage data, trap data were used to infer fruit fly activity at trapping locations only.

Table 2. The total number of viable *Ceratitis capitata* traps per crop type, used for analysis, in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa.

Region	Crop type	Number of traps
Elgin/Grabouw	Grapes	8
	Home gardens	16
	Pome fruit	228
	Stone fruit	7
Villiersdorp	Citrus	1
	Pome fruit	50
	Stone fruit	17
Vyeboom	Berries (Mulberry)	1
	Citrus	2
	Pome fruit	66
	Stone fruit	3

Trap catches are presented as the total number of flies (males and females combined) per trap per week (FTW) for each sub-region of EGVV. The mean FTW was calculated for each

season per region by dividing the total number of flies captured per season by the product of the number of serviced traps and the number of weeks that the traps were active in the field. Regions and seasons were compared and the data (mean FTW) were compared by means of the nonparametric Kruskal-Wallis test (using SPSS version 25.0 (IBM corp 2017)) because FTW values were not normally distributed (Shapiro-Wilk test: $p < 0.05$). To determine significant differences between groups, a Dunn's post hoc test with Bonferroni corrections was carried out.

For a visual representation of the spatio-temporal distribution of *C. capitata*, trap catch data (mean FTW) were arranged by column with a unique identifier (trap name) arranged by row in Microsoft Excel. Data were then imported into a GIS database using the ArcGIS 10.5® desktop package (ESRI 2016). The x and y coordinates of traps, with the trap name as unique identifier, were imported and converted from a Google Earth KML layer to a feature class into the same GIS database. Using the unique identifier, trap catch data imported from Microsoft Excel were joined to the trap location feature class and saved in the GIS database. All feature class data in the GIS database were projected to the UTM Zone 34S coordinate system (WGS 1984 reference surface/datum) to allow for analyses. Seasonal and weekly FTW maps were created for each season and for all seasons combined. All the maps and feature class data were stored in a GIS database.

The trap catches were classified and symbolised in the GIS to support visual interpretation of the maps. The classifications were based on subjective population thresholds used because no action thresholds have been statistically determined to guide management actions. These subjective population thresholds are based on the number of flies per trap per day (FTD). Barnes *et al.* (2015) indicated that an FTD of up to 0.5 indicate a low population pressure, while an FTD between 0.6 and 1.0 indicates a moderate population

pressure and an FTD of more than one indicates that the population pressure is high in an area. In this study, this subjective FTD-threshold was extrapolated to FTW, as a week number and not a date was entered in the records. For the extrapolation we assumed that one week consisted of seven days. Therefore, a threshold of seven FTW was used to indicate high population pressure, which relates to one FTD. This was done to standardise data and facilitate comparisons.

Seasonal weather data used in this study were obtained from two separate automatic weather stations located in the Vyeboom (-34,04088° 19,14193° - 308 m) and Grabouw (-34,16517° 19,03155° - 300 m) regions. No reliable weather station was available in Villiersdorp. However, the Vyeboom weather station is close (±10 km) to Villiersdorp and therefore was used to represent Villiersdorp and Vyeboom. Weather data were supplied by the Agricultural Research Council of South Africa.

Results

The results showed that the mean trap catch of *C. capitata* did not significantly differ between the three sub-regions of EGVV (p = 0.132) (Figure 5).

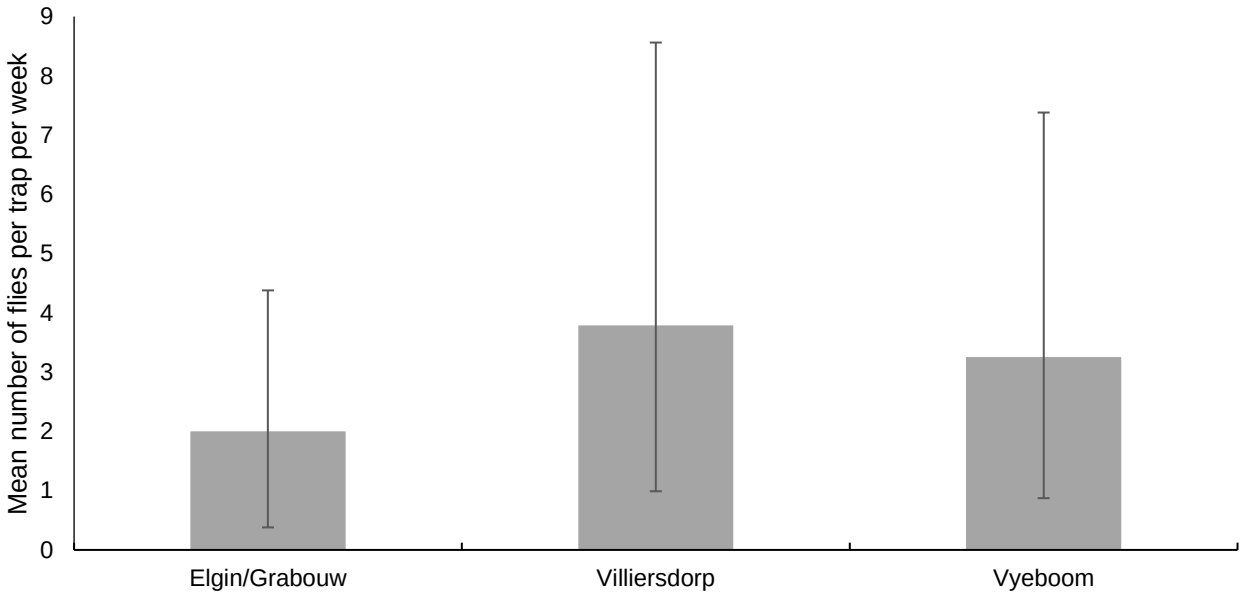

Figure 5. Mean *Ceratitis capitata* trap catch per week in Elgin/Grabouw, Villiersdorp and Vyeboom over four consecutive seasons (2010/2011 – 2013/2014). Error bars denote ± 95% confidence intervals.

Ceratitis capitata trap catches varied between seasons within the different regions (Figure 6). In Elgin/Grabouw there were no significant differences between seasons (p = 0.089), while significant differences between 2010/2011 and 2012/2013 (p < 0.001), as well as between 2010/2011 and 2013/2014 (p < 0.05), were observed in Villiersdorp. Likewise, significant differences were recorded for Vyeboom between 2010/2011 and all subsequent seasons (p < 0.05), but not between 2011/2012 and 2012/2013, 2013/2014 and between 2012/2013 and 2013/2014 (p = 1.000).

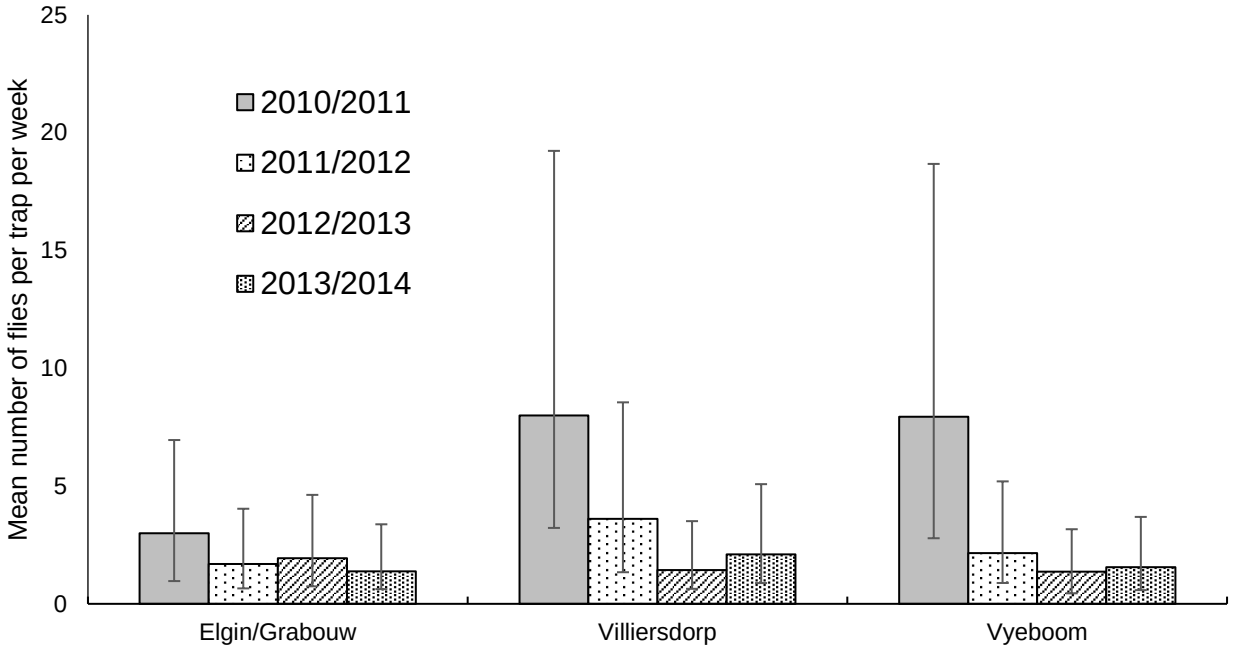

Figure 6. Mean seasonal *Ceratitis capitata* trap catch per week for seasons 2010/2011 to 2013/2014 in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa. Error bars denote ± 95% confidence intervals.

The spatial distribution of the mean seasonal *C. capitata* trap catch per week over all seasons indicated no definitive spatial distribution pattern, although traps with a mean seasonal FTW of more than seven occurred mostly in the south-eastern parts of each region and traps with lower levels were distributed mostly in the north-western parts of each region (Figure 7).

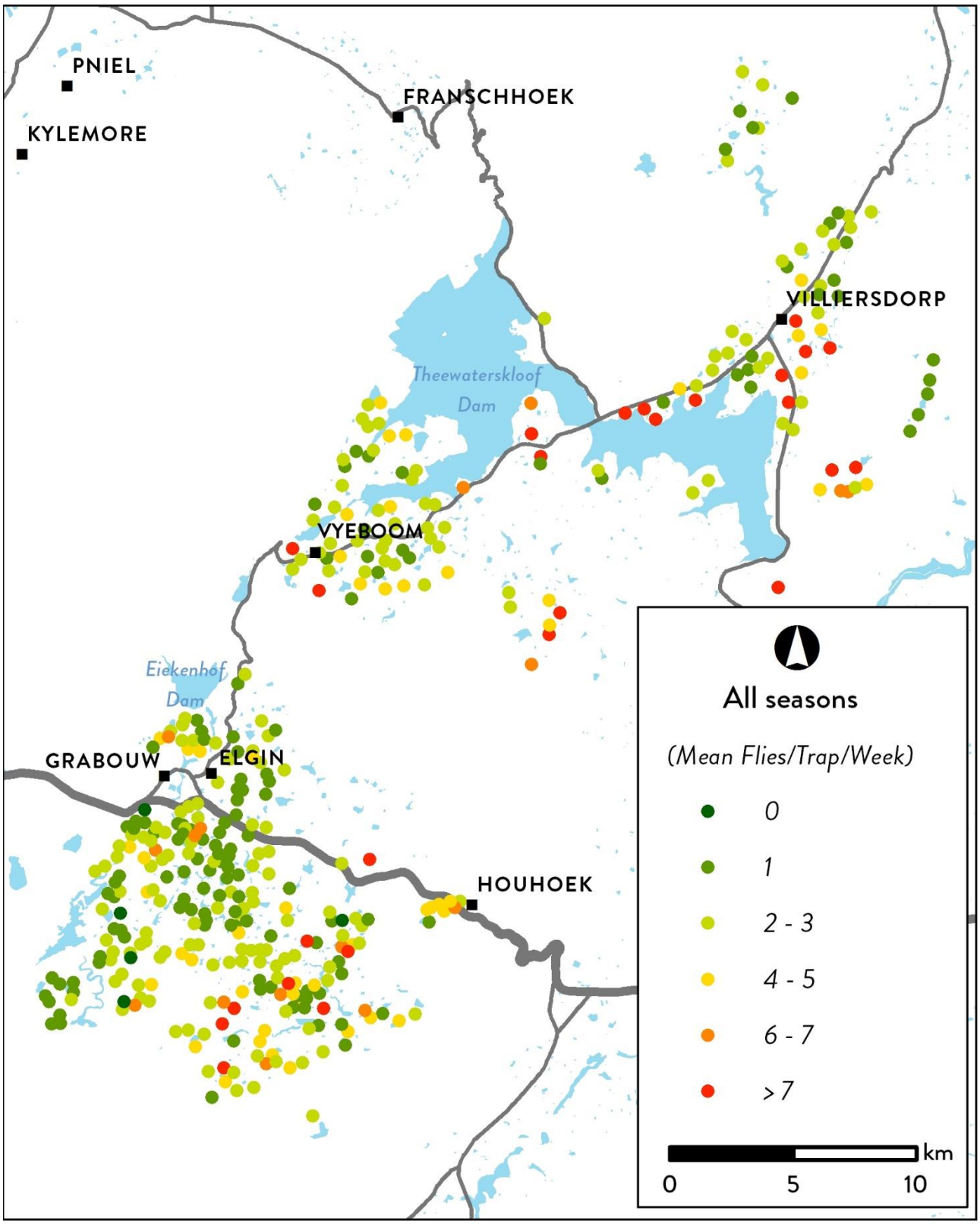

Figure 7. Spatial distribution of the mean seasonal *Ceratitis capitata* trap catch per week for all seasons combined (2010/2011 – 2013/2014) in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa.

The spatial distribution of the mean seasonal trap catch per week varied between seasons (Figure 8a–d). A distinct decrease in the number of traps with a mean seasonal trap catch (FTW) of more than seven was observed from 2010/2011 (Figure 8a) to 2013/2014 (Figure 8d). The spatial distribution was highly variable, but consistently high levels of trap catches were noted in some smaller areas.

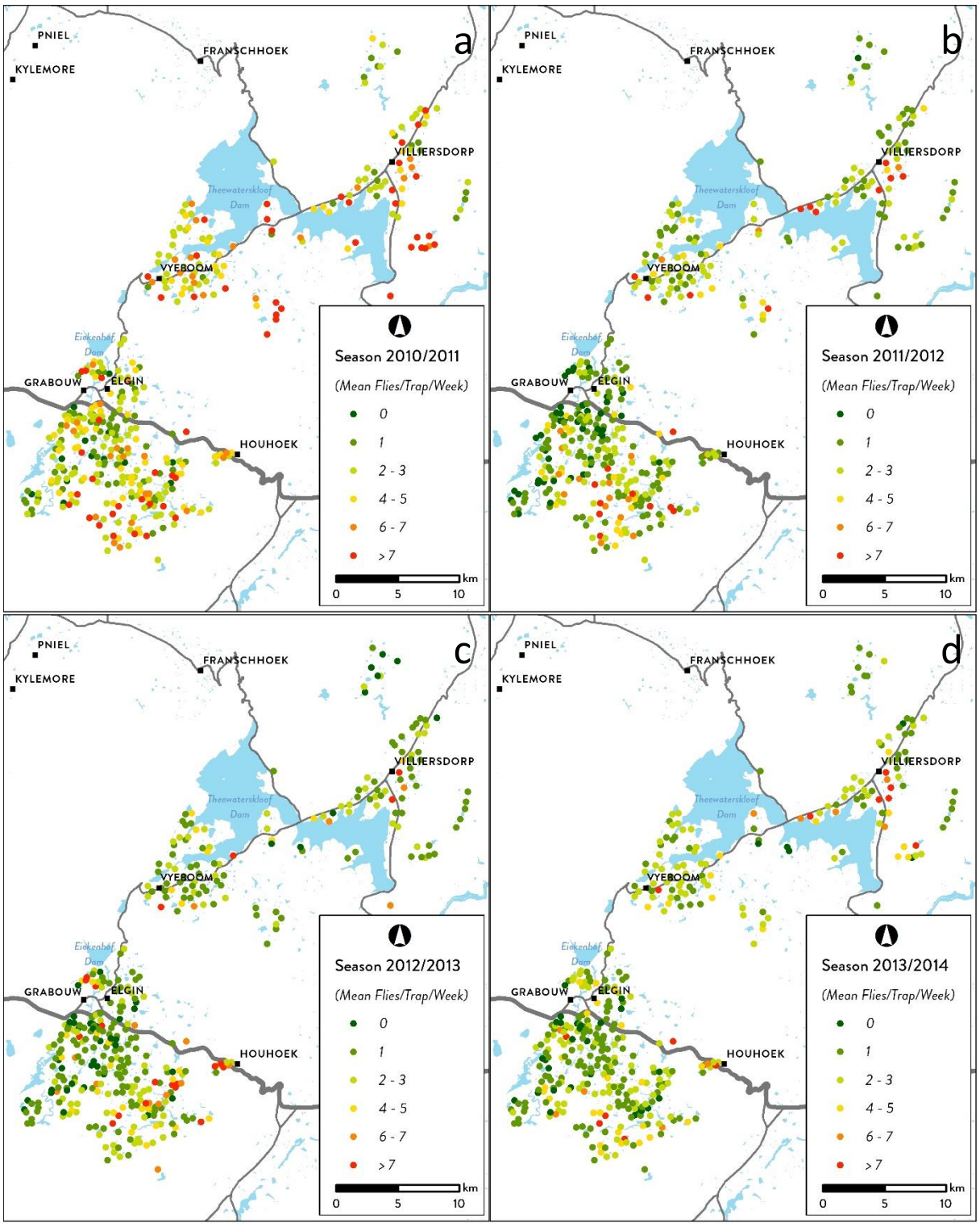

Figure 8. Spatial distribution of the mean seasonal *Ceratitis capitata* trap catch per week for the a) 2010/2011, b) 2011/2012, c) 2012/2013 and d) 2013/2014 season in Elgin/Grabouw, Villiersdorp and Vyeboom Western Cape, South Africa.

The spatial distribution of *C. capitata* trap catch during peak weeks also varied between seasons (Figure 9a–d), with some areas consistently yielding high trap catches, indicating population hot spots.

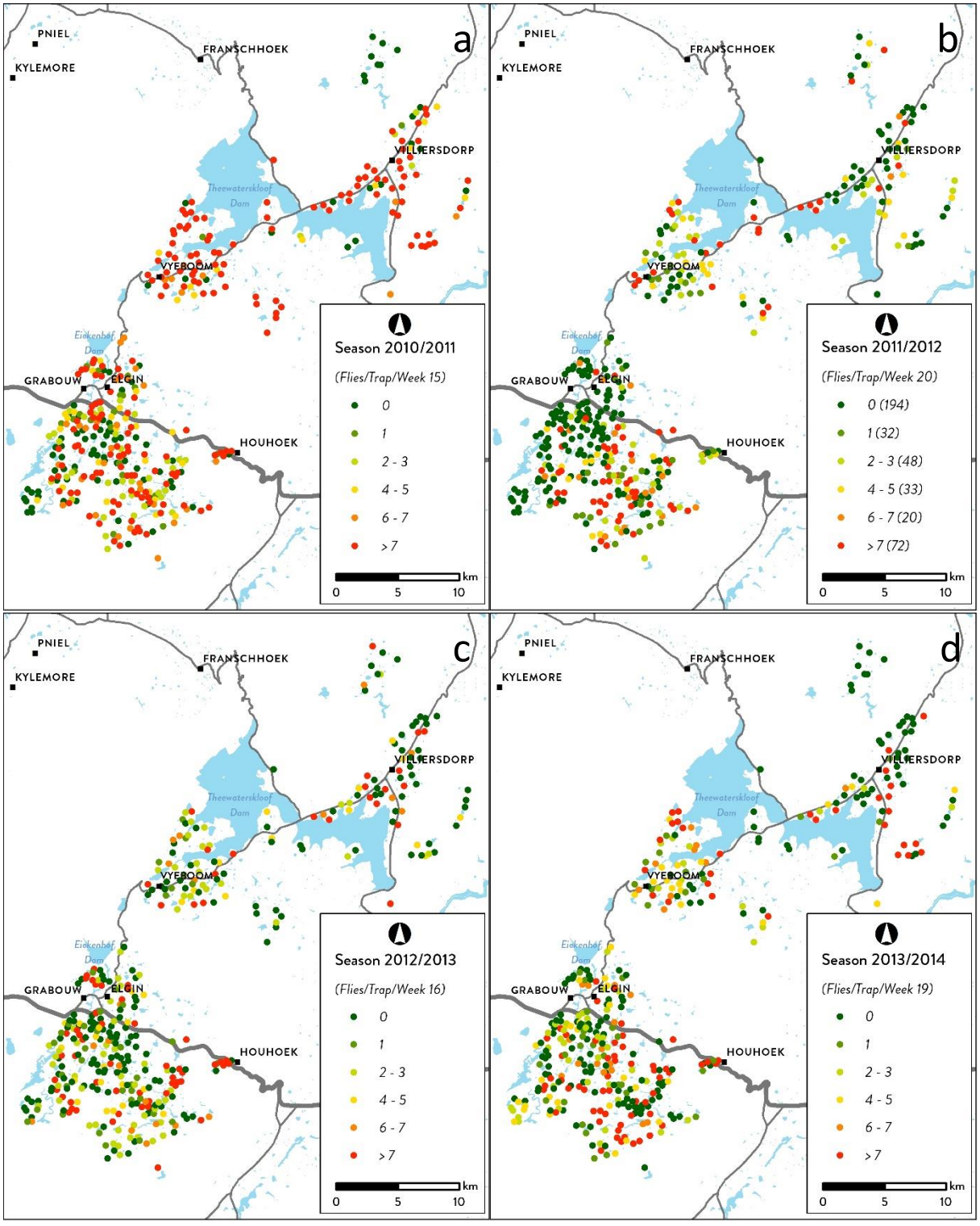

Figure 9. The spatial distribution of *Ceratitis capitata* trap catch per week for weeks when fruit fly populations peaked in Elgin/Grabouw, Villiersdorp and Vyeboom in each season a) 2010/2011, b) 2011/2012, c) 2012/2013 and d) 2013/2014.

Weekly trapping activity was low, even negligible, for the first part of the season (November to December) in all regions and all seasons (Figure 10a–d). Peaks in catches were reached at the end of each season (March to May) (Figure 10a–d), but the timing of when these peaks occurred varied between seasons. Maximum population peaks occurred from April to May, except for 2011/2012 where the maximum population peak in the Villiersdorp region occurred in March, and an additional early peak was recorded in Vyeboom during February in 2010/2011.

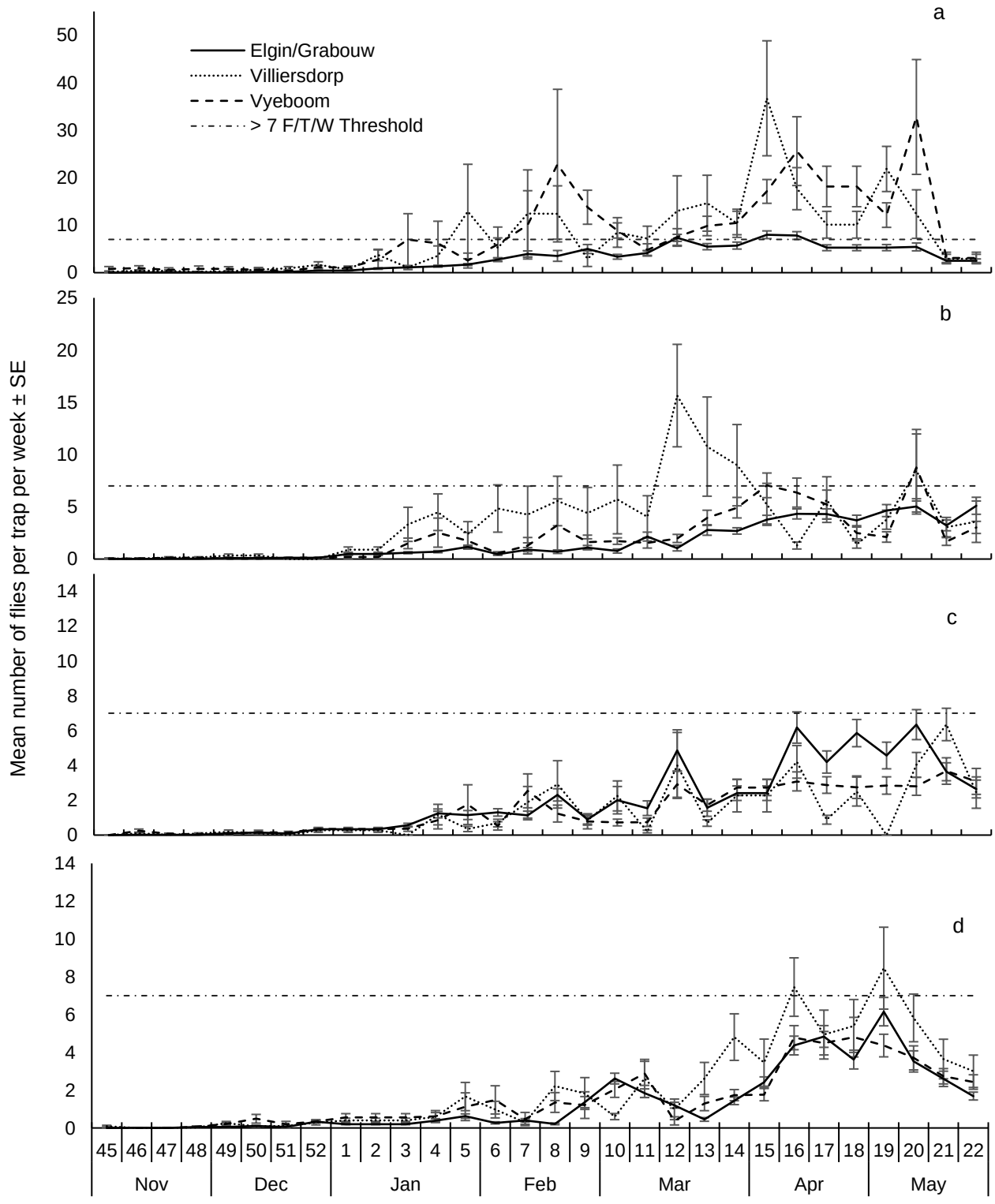

Figure 10. Mean weakly (± SE) *Ceratitis capitata* trap catch fluctuations indicated the by the x-axis (week 45 to 22) per season a) 2010/2011, b) 2011/2012, c) 2012/2013 and d) 2013/2014, in Elgin/Grabouw, Villiersdorp and Vyeboom, Western Cape, South Africa. The threshold line (---) indicates a mean FTW of seven.

Winter/spring (June to November) weather patterns, varied among the different years and between the regions under study (Table 3). The total rainfall received between the years within each region gradually increased from 2010 to 2013 during the winter/spring period in EGVV, while there was minimal variation in the minimum (0.22 – 0.28 standard errors) and maximum temperatures (0.33 – 0.38 standard errors) (Table 3).

Table 3. Mean (±SE) maximum, minimum temperatures (°C) and total rainfall (mm) for the winter/spring period (June to November) for the Vyeboom and Villiersdorp regions combined and the Elgin/Grabouw region, from the year 2010 to 2013.

Region	Year	Mean (±SE) min. temp. (°C)	Mean (±SE) max. temp. (°C)	Total Rainfall (mm)
Vyeboom/Villiersdorp	2010	9,39 ± 0,26	20 ± 0,34	314,20
	2011	9,96 ± 0,22	20,28 ± 0,35	480,50
	2012	8,19 ± 0,25	18,23 ± 0,36	621,10
	2013	9,09 ± 0,25	18,9 ± 0,38	798,30
Elgin/Grabouw	2010	7,43 ± 0,27	19,42 ± 0,33	471,80
	2011	7,12 ± 0,24	18,7 ± 0,33	474,12
	2012	7,22 ± 0,25	17,82 ± 0,33	618,65
	2013	7,39 ± 0,28	18,02 ± 0,34	1058,77

Discussion

Climatic differences between locations, as well as host availability and suitability, have been shown to influence the abundance and distribution of *C. capitata* populations (Nishida *et al.* 1985; Segura *et al.* 2006). *Ceratitis capitata* prefers hot and dry areas over wet and cold areas (Israely *et al.* 2005b; Duyck *et al.* 2006b). The warmer and drier climate of Vyeboom and Villiersdorp (compared to Elgin/Grabouw) might thus have favoured the development and survival of *C. capitata* populations and could have contributed to the relatively higher mean trap catches in these areas. In terms of host availability and suitability, stone fruit (e.g.

peaches, plums and nectarines) and pome fruit (e.g. pears) are highly suitable hosts for *C. capitata* (Nishida *et al.* 1985; Krainacker *et al.* 1987; DeVilliers *et al.* 2013; Manrakhan & Addison 2014). Furthermore, Papadopoulos *et al.* (2002) reported high larval mortality and long larval development periods for *C. capitata* reared on apples. These development times were longer compared to the larval development times for softer-skinned stone fruit (Carey 1984). Villiersdorp has the largest proportion of stone fruit and is consequently more prone to fruit fly attacks than Elgin/Grabouw, where mainly apples and a relatively small amount of stone fruit are planted. This indicates that fruit fly populations are influenced differently by several factors, justifying a more in-depth analysis of the spatio-temporal distribution patterns associated particularly with host plant and climate variables. For example, there appears to be a spatial trend with higher trap catches occurring in the south-eastern parts of each region. The consistency of this trend suggests that the distribution is most likely due to climatic differences.

The significant differences of trap catches between the seasons within each region, as well as the decreasing trend in trap catches over time are likely due to climatic variations and management differences. Low temperatures have been shown to limit the survival and development of *C. capitata* populations, while warmer climates extend the season over which the flies are active and able to reproduce (Nyamukondiwa *et al.* 2013). Rainfall and temperature, as well as the interaction between the two, have been found to affect the abundance of *C. capitata* populations (Duyck *et al.* 2006b). Higher rainfall and lower temperatures during the winter/spring period negatively influence fruit fly development and survival, leading to lower seasonal populations (Escudero-Colomar *et al.* 2008; Peñarrubia-María *et al.* 2012b; Flores *et al.* 2016). This is also true for the current findings, where increased rainfall during the winter/spring period, from 2010 to 2013, coincides with a decrease in mean trap catches, from 2010/2011 to 2013/2014, in EGVV. Rainfall in the

winter/spring period seems to have played a more prominent role than temperature in influencing seasonal fruit fly populations, as there was minimal variability in the winter/spring minimum and maximum temperatures between the different years. The between-season differences in the mean trap catch within the sub-regions of EGVV were also likely affected by the differences in fruit fly management efforts, as the number of aerial BAT applications was doubled from 2011/2012 (four applications) onward, compared to the 2010/2011 season (two applications). This likely contributed to explaining the significant decrease in the mean trap catch observed from 2010/2011 to 2011/2012 in Vyeboom. Although Vyeboom was the only region where the decrease in mean trap catch between 2010/2011 and 2011/2012 was significant, there were clear decreases in the mean trap catch from 2010/2011 to 2011/2012 in the other two regions. Apart from area-wide management activities by FFA, individual growers also apply orchard management, which differs among regions, individual growers and seasons. It is possible that these management activities also affected the mean trap catch between the different regions and the seasons, but the extent of these activities is difficult to gauge. However, it is likely that the impact of these activities would be negligible at regional scales.

Winter/spring temperatures have been shown to affect the timing of the population peaks within the season (Escudero-Colomar *et al.* 2008; Peñarrubia-María *et al.* 2012b). This is not only due to the effect temperature has on the developmental time of *C. capitata* (Duyck & Quilici 2002; Nyamukondiwa *et al.* 2013), but the timing of the population peak is also affected by the ripening of fruit (e.g. become susceptible to fruit fly attack) in the summer and autumn and the length of the growing season (Menzel & Sparks 2006). In eastern Spain, lower average daily temperatures in the winter/spring period resulted in *C. capitata* populations peaking later in the season compared to when winter/spring temperatures where higher (Martinez-Ferrer *et al.* 2010). Therefore, the maximum population peak that

occurred in March of the 2011/2012 season was most likely due to the relatively higher winter/spring temperatures of 2011 in the Villiersdorp/Vyeboom region, resulting in earlier and faster development of the winter/spring population.

Conclusion

Through the descriptive analysis of the area-wide trap catches and the visual representation of the spatio-temporal distribution of the mean *C. capitata* trap catches, it is clear that these populations are highly variable, spatially and temporally. The results suggest a relationship between the geographic characteristics of EGVV and the abundance (and distribution) of *C. capitata*. However, the high variability in the spatial distribution of the data (per trap location) limits the identification of patterns from which management-supporting recommendations can be made. Given that management decisions in AW-IPM programmes are generally taken at coarser spatial scales, it is recommended that the data are spatially and temporally aggregated into manageable units of about 1 km^2. This will allow for improved statistical quantifications and likely accentuate the spatial relationships between the fruit fly distributions and the geographic characteristics within the area (Midgarden *et al.* 2014).

Rainfall, temperature, host availability and host suitability seem to have played a role in determining fruit fly activity in EGVV. However, it is not yet clear to what extent each of these factors contributed, and quantification of the impact of each of these factors is needed. Data mining techniques and geospatial analyses may help to identify the main underlying factors that drive fruit fly spatial distributions in large heterogeneous fruit production areas like EGVV. The spatial relationships between the different environmental factors, management interventions and distribution of trap catches must be investigated. Specifically, identifying areas of significance (based on trap catches), and investigating the relationship between these significant areas and geographic characteristics is critical. This information will be of

great value to AW-IPM programme managers as it will support management decisions and refine the application of management actions against fruit flies on an area-wide scale.

References

ALEMANY, A., MIRANDA, M.A., ALONSO, R. & MARTÍN ESCORZA, C. 2006. Changes in the spatial and temporal population density of the Mediterranean fruit fly (Diptera: Tephritidae) in a citrus orchard. *Spanish Journal of Agricultural Research* **4**(2): 161–166.

BARNES, B.N. 2009. The Sterile Insect Technique (S.I.T.) for management of Mediterranean Fruit Flies in fruit production - A technical manual. Stellenbosch, South Africa.

BARNES, B.N., HOFMEYR, J.H., GROENEWALD, S., CONLONG, D.E. & WOHLFARTER, M. 2015. The sterile insect technique in agricultural crops in South Africa : a metamorphosis .... but will it fly ? *African Entomology* **23**(1): 1–18.

BARNES, B.N., ROSENBERG, S., ARNOLDS, L. & JOHNSON, J. 2007. Production and quality assurance in the SIT Africa Mediterranean Fruit Fly (Diptera: Tephritidae) rearing facility in South Africa. *Florida Entomologist* **90**(1): 41–52.

BARNES, B.N. & VENTER, J.H. 2006. The South African Fruit Fly Action Plan – Area-wide Suppression and Exotic Species Surveillance. *Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*. 271–283. Salvador, Brazil.

BECKLER, A.A., FRENCH, B.W. & CHANDLER, L.D. 2005. Using GIS in Areawide Pest Management: A Case Study in South Dakota. *Transactions in GIS* **9**(2): 109–127.

BERRYMAN, A.A. 2003. On Principles, Laws and Theory in Population Ecology. *Oikos* **103**(3): 695–701.

CAREY, J.R. 1984. Host-specific demographic studies of the Mediterranean fruit fly *Ceratitis capitata*. *Ecological Entomology* **9**(3): 261–270.

CASTRIGNANÒ, A., BOCCACCIO, L., COHEN, Y., NESTEL, D., KOUNATIDIS, I., PAPADOPOULOS, N.T., DE BENEDETTO, D. & MAVRAGANI-TSIPIDOU, P. 2012. Spatio-temporal population dynamics and area-wide delineation of *Bactrocera oleae* monitoring zones using multi-variate geostatistics. *Precision Agriculture* **13**(4): 421–441.

COX, J.S.H. & VREYSEN, M.J.B. 2005. Use of Geographic Information Systems and Spatial Analysis in Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 453–477. Springer, Dordrecht.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DMINIC, I., KOZINA, A., BAZOK, R. & BARCIC, J.I. 2010. Geographic information systems (GIS) and entomological research: A review. *Journal of Food, Agriculture & Environment* **8**(2): 1193–1198.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

DUYCK, P.F. & QUILICI, S. 2002. Survival and development of different life stages of three *Ceratitis* spp. (Diptera : Tephritidae) reared at five constant temperatures. *Bulletin of Entomological Research* **92**(6): 461–469.

EKESI, S., LUX, S.A. & BILLAH, M. 2005. Field comparison of food-based synthetic attractants and traps for African Tephritid fruit flies. *Development of Improved Attractants and Their Integration into Fruit Fly SIT Management Programmes*. 205–230. International Atomic Energy Agency, Vienna.

ENKERLIN, W., VILLASEÑOR, A., FLORES, S., MIDGARDEN, D., LIRA, E., RENDON, P., HURLEY, J., SALAZAR, E., MENDEZ, W., CASTAÑEDA, R., COTOC, E., ZAVALA, J.L., CELEDONIO, H. & GUTIERREZ-RUELAS, J.M. 2016. Descriptive analysis of the factors affecting the population dynamics of the Mediterranean fruit fly (*Ceratitis capitata*, Wied.) in coffee areas located in Guatemala and its implications in IPM Strategies. In: Sabater-Muñoz, B., Vera, M.T., Pereira, R. & Orankanok, W. (Eds.) *Proceedings IX International Symposium on Fruit Flies of Economic Importance, 12-14 May 2014*. 46–63.

EPSKY, N.D., ESPINOZA, H.R., KENDRA, P.E., ABERNATHY, R., MIDGARDEN, D. & HEATH, R.R. 2010. Effective Sampling Range of a Synthetic Protein-Based Attractant for *Ceratitis capitata* (Diptera: Tephritidae). *Journal of Economic Entomology* **103**(5): 1886–1895.

EPSKY, N.D., HENDRICHS, J., KATSOYANNOS, B.I., VASQUEZ, L.A., ROS, J.P., ZUMREOGLU, A., PEREIRA, R., BAKRI, A., SEEWOORUTHUN, S.I. & HEATH, R.R. 1999. Field Evaluation of Female-Targeted Trapping Systems for *Ceratitis capitata* (Diptera:Tephritidae) in Seven Countries. *Journal of Economic Entomology* **92**: 156–164.

ESCUDERO-COLOMAR, L.A., VILAJELIU, M. & BATLLORI, L. 2008. Seasonality in the occurrence of the Mediterranean fruit fly [*Ceratitis capitata* (Wied.)] in the north-east of Spain. *Journal of Applied Entomology* **132**(9–10): 714–721.

ESRI. 2016. ArcGIS Release 10.5. Redlands, CA.

FLORES, S., MONTOYA, P., RUIZ-MONTOYA, L., VILLASEÑOR, A., VALLE, A., ENKERLIN, W. & LIEDO, P. 2016. Population Fluctuation of *Ceratitis capitata* (Diptera: Tephritidae) as a Function of Altitude in Eastern Guatemala. *Environmental Entomology* **45**(4): 802–811.

GUIDOTTI, D., RAGAGLINI, G. & PETACCHI, R. 2005. Analysis of spatio-temporal *Bactrocera oleae* (Diptera, Tephritidae) infestation distributions obtained from a large-scale monitoring network and its importance to IPM. *Integrated Protection of Olive Crops IOBC/wprs Bulletin* **28**(9): 13–18.

HEATH, R.R., EPSKY, N.D., DUEBEN, B.D., RIZZO, J. & JERONIMO, F. 1997. Adding methyl-substituted ammonia derivatives to a food-based synthetic attractant on capture of the Mediterranean and Mexican fruit flies (Diptera: tephritidae). *Journal of Economic Entomology* **90**(6): 1584–1589.

HENDRICHS, J., KENMORE, P. & ROBINSON, A.S. 2007. Area-wide integrated pest management (AW-IPM): principles, practice and prospects. In: Vreysen, M.J.B., Robinson, A.S. & Hendrichs, J. (Eds.) *Area-wide control insect pests: from research to field implementation*. 3–33. Springer, Dordrecht.

HENDRICHS, J., REYES, J., VREYSEN, M.J.B. & CARDOSO-PEREIRA, R. 2015. Area-Wide Management of Major Pest Insects. In: Sabater-Munoz, B., Pena, L., Navarro, L. & Moreno, P. (Eds.) *XII International Citrus Congress - International Society of Citriculture*. 67–72. Acta Horticulture, Valencia (Spain).

IAEA. 2003. Trapping guidelines for area-wide fruit fly programmes. International Atomic Energy Agency, Vienna.

IBM CORP. 2017. IBM SPSS Statistics for Windows, Version 25. Armonk, NY: IBM Corp.

ISRAELY, N., ZIV, Y. & GALUN, R. 2005a. Metapopulation Spatial – Temporal Distribution Patterns of Mediterranean Fruit Fly (Diptera: Tephritidae) in a Patchy Environment. *Annals of the Entomological Society of America* **98**(3): 302–308.

ISRAELY, N., ZIV, Y. & OMAN, S.D. 2005b. Spatiotemporal Distribution Patterns of Mediterranean Fruit Fly (Diptera : Tephritidae) in the Central Region of Israel. *Annals of the Entomological Society of America* **98**(1): 77–84.

KARSTEN, M. 2014. Population genetic stucture and abundance of two *Ceratitis* species (Tephritidae) of agricultural importance in South Africa. Stellenbosch University.

KITRON, U., OTIENO, L.H., HUNGERFORD, L.L., ODULAJA, A., BRIGHAM, W.U., OKELLO, O.O., JOSELYN, M., MOHAMED-AHMED, M.M. & COOK, E. 1996. Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS. *Journal of Animal Ecology* **65**(3): 371–380.

KOUNATIDIS, I., PAPADOPOULOS, N.T., MAVRAGANI-TSIPIDOU, P., COHEN, Y., TERTIVANIDIS, K., NOMIKOU, M. & NESTEL, D. 2008. Effect of elevation on spatio-temporal patterns of olive fly (*Bactrocera oleae*) populations in northern Greece. *Journal of Applied Entomology* **132**: 722–733.

KRAINACKER, D.A., CAREY, J.R. & VARGAS, R.I. 1987. Effect of larval host on life history traits of the mediterranean fruit fly, *Ceratitis capitata*. *Oecologia* **73**(4): 583–590.

LEZA, M.M., JUAN, A., CAPLLONCH, M. & ALEMANY, A. 2008. Female-biased mass trapping vs . bait application techniques against the Mediterranean fruit fly, *Ceratitis capitata* (Dipt., Tephritidae). *Journal of Applied Entomology* **132**(9–10): 753–761.

LINDQUIST, D.A. 2000. Pest Management Strategies: Area-wide and Conventional. In: Tan, K. (Ed.) *Area-wide control of fruit flies and other insect pests. Joint proceedings of the international conference on area-wide control of insect pests and the fifth international symposium on fruit flies of economic importance*. 13–19. Penerbit Universiti Sains Malaysia, Penang.

MANRAKHAN, A. & ADDISON, P. 2014. Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science* **70**(4): 651–660.

MARTINEZ-FERRER, M.T., NAVARRO, C., CAMPOS, J.M., MARZAL, C., FIBLA, J.M., BARGUES, L. & GARCIA-MARI, F. 2010. Seasonal and annual trends in field populations of Mediterranean fruit fly, *Ceratitis capitata*, in Mediterranean citrus groves: comparison of two geographic areas in eastern Spain. *Spanish Journal of Agricultural Research* **8**(3): 757–765.

MCGOVERN, T., RENDON, P. & LIRA, E. 2008. Programme operational update: Targeted suppression of population centers of *Ceratitis capitata* Wied through the use of GIS/GPS technology. In: Fleischer, F.D., Gerardo, P.J.M. & Breceda, S.F. (Eds.) *Proceedings of the 7th Meeting of the Working Group on Fruit Flies of the Western Hemisphere*. 9–10. Mexico.

MENZEL, A. & SPARKS, T. 2006. Temperature and plant development: phenology and seosonality. In: Morecroft, M.D. & Paterson, J.S. (Eds.) *Plant growth and climate change*. 70–93. Blackwell Publishing.

MIDGARDEN, D. & LIRA, E. 2006. Ecological relationship of Medfly and coffee in guatemala and Mexico. *Fruit Flies of Economic Importance: From Basic to Applied Knowledge Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*: 241–247.

MIDGARDEN, D.G., LIRA, E. & SILVER, M. 2014. Spatial Analysis of Tephritid Fruit Fly Traps. In: Shelly, T., Epsky, N., Jang, E.B., Reyes-Floris, J. & Vargas, R. (Eds.) *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. 277–320. Springer, Dordrecht.

MIRANDA, M.A., ALONSO, R. & ALEMANY, A. 2001. Field evaluation of Medfly (Dipt., Tephritidae) female attractants in a Mediterranean agrosystem (Balearic Islands, Spain). *Journal of Applied Entomology* **125**(6): 333–339.

MUMFORD, J. & TWEEN, G. 1997. Economic feasibility study for the control of the Mediterranean and Natal fruit fly in the Western Cape Province. Expert Mission Report RU-7135. Vienna.

MYBURGH, A.C. 1956. Bionomics and control of the fruit flies, *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.), in the Western Cape. University of Stellenbosch.

MYERS, N., MITTERMEIER, R.A., MITTERMEIER, C.G., DA FONSECA, G.A.B. & KENT, J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**(6772): 853–858.

NESTEL, D., CARVALHO, L. & NEMNY-LAVY, E. 2004. The spatial dimension in the ecology of insect pests and its relevance to pest management. In: Horowitz, A.R. & Ishaaya, I. (Eds.) *Insect Pest Management*. 45–63. Springer, Heidelberg, Berlin.

NISHIDA, T., HARRIS, E.J., VARGAS, R.I. & WONG, T.T.Y. 1985. Distributional loci and host fruit utilization patterns of the Mediterranean fruit fly, *Ceratitis capitata* (Diptera: Tephritidae), in Hawaii. *Environmental Entomology* **14**(5): 602–606.

NYAMUKONDIWA, C., WELDON, C.W., CHOWN, S.L., LE ROUX, P.C. & TERBLANCHE, J.S. 2013. Thermal biology, population fluctuations and implications of temperature extremes for the management of two globally significant insect pests. *Journal of Insect Physiology* **59**(12): 1199–1211.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & CAREY, J.R. 2002. Demographic Parameters of the Mediterranean Fruit Fly (Diptera: Tephritidae) Reared in Apples. *Annals of the Entomological Society of America* **95**(5): 564–569.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PEÑARRUBIA-MARÍA, E., AVILLA, J., ESCUDERO-COLOMAR, L.A., PENAEEUBIA-MARIA, E., AVILLA, J. & ESCUDERO-COLOMAR, L.A. 2012. Survival of wild adults of *Ceratitis capitata* (Wiedemann) under natural winter conditions in North East Spain. *Psyche* **2012**: 1–6.

ROESSLER, Y. 1989. Insecticidal bait and cover sprays. In: Robinson, A.S. & Hooper, G. (Eds.) *World Crop Pests: Fruit Flies - Their Biology, Natural Enemies and Control*. 329–335. Elsevier, Amsterdam.

SCHULZE, R. 2006. *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

SCIARRETTA, A. & TREMATERRA, P. 2014. Geostatistical Tools for the Study of Insect Spatial Distribution: Practical Implications in the Integrated Management of Orchard and Vineyard Pests. *Plant Protection Science* **50**(2): 97–110.

SEGURA, D.F., VERA, M.T., CAGNOTTI, C.L., VACCARO, N., DE COLL, O., OVRUSKI, S.M. & CLADERA, J.L. 2006. Relative Abundance of *Ceratitis capitata* and *Anastrepha fraterculus* (Diptera: Tephritidae) in Diverse Host Species and Localities of Argentina. *Annals of the Entomological Society of America* **99**: 70–83.

VARGAS, R.I., HARRIS, E.J. & NISHIDA, T. 1983. Distribution and Seasonal Occurrence of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) on the Island of Kauai in the Hawaiian Islands. *Environmental Entomology* **12**(2): 303–310.

WESTERN CAPE DEPARTMENT OF AGRICULTURE. 2014. The Mapping of Agricultural Commodity Production and Agricultural Infrastructure in the Western Cape.

WHITE, I.M. & ELSON-HARRIS, M.M. 1992. *Fruit flies of economic significance: their identification and bionomics*. CAB International/Australian Centre for International Agricultural Research, Wallingford, UK/Canberra.

WONG, T.T.Y., NISHIMOTO, J.I. & MOCHIZUKI, N. 1983. Infestation Patterns of Mediterranean Fruit Fly and the Oriental Fruit Fly (Diptera: Tephritidae) in the Kula Area of Maui, Hawaii. *Environmental Entomology* **12**(4): 1031–1039.

Chapter 4

Using machine learning to identify the geographical drivers of *Ceratitis capitata* trap catch in an agricultural landscape

This chapter has been accepted with minor revisions in *Computers and Electronics in Agriculture* as Bekker, G.F.H.v.G., Addison, M.F., Addison, P. and Van Niekerk, A. Using machine learning to identify the geographical drivers of *Ceratitis capitata* trap catch in an agricultural landscape

Abstract

The spatial distribution of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) trap catch was classified and related to a set of geographic variables, to identify its main geographical drivers. Trap catch data were sourced from an area-wide integrated pest management (AW-IPM) programme and classified into statistically significant hot- and cold spots (HCSs). Trap data of four consecutive fruiting seasons were combined to identify monthly and seasonal long-term HCSs. The main geographic drivers of the HCSs were identified using variable importance lists (VIL's) produced by the random forest (RF) machine learning (ML) algorithm. Long-term climate, topography, landscape and fruit fly management variables were used as predictor variables in RF to classify HCSs. The resulting RF models produced classification accuracies of up to 80%. In most cases, the most important variable was long-term rainfall, suggesting that this was the most prominent driver of *C. capitata* HCSs in our study region. The result of this study highlights the value of long-term pest monitoring data and, long-term environmental data in improving our understanding of the spatial distribution of *C. capitata* trap catch in complex agricultural systems. This study sets out a framework to spatially quantify *C. capitata* trap catch into HCSs using monitoring data from area-wide integrated pest management (AW-IPM) programmes, enabling the investigation of complex ecological relationships through the use of ML algorithms. The information gained could be used in area-wide integrated fruit fly management programmes, which incorporates the sterile insect technique (SIT), to conduct more precise spatial planning, leading to better programme performance and reduced costs.

Introduction

Tephritid fruit flies (Diptera: Tephritidae) are known to attack fruits and flowers. About 250 of the 4000 known species attack a variety of commercially grown fruits globally (White & Elson-Harris 1992), causing excessive economic losses to horticultural industries in many regions of the world. In order to implement management or control actions regarding these flies, it is important to know and understand the spatial distribution of these pests (Papadopoulos *et al.* 2003; Hendrichs *et al.* 2007a). However, the response of insects, including fruit flies, to the changing environment (abiotic factors) determines their geographical distribution (Vargas *et al.* 1983; Nestel *et al.* 2002; Papadopoulos *et al.* 2003; Israely *et al.* 2005a; Israely *et al.* 2005b; Lindsay 2010; Mazzi & Dorn 2012). Furthermore, the geographical distribution of species is also driven by biotic and/or ecological processes, which are inherently spatial (Hendrichs & Hendrichs 1990; Wagner & Fortin 2005; Hortal *et al.* 2010; Mazzi & Dorn 2012; Sciarretta & Trematerra 2014). These underlying geographic drivers or processes (biotic and abiotic factors) interact with one another across multiple spatial scales, and therefore the spatial response of species to these processes changes as the scale of the observation is altered (Whittaker *et al.* 2001; Soberón & Nakamura 2009). Thus, the determinants of a species' spatial distribution is complex and cannot be explained by a single process, but are rather the result of the species' response to multiple processes at different spatio-temporal scales (Wagner & Fortin 2005). This poses a challenge to explaining, predicting and responding to the spatial distribution of a species.

Geospatial technologies – e.g. geographic position systems (GPS), geographic information systems (GIS) and spatial- or geostatistics, have been shown to be invaluable for studying the spatial distribution of pests (Nestel *et al.* 2002; Soberón 2007; Gutierrez & Ponti 2011). These methods are also useful for identifying areas where large pest populations persist and for finding spatial distribution patterns (Cox & Vreysen 2005; Abd El Wahab *et al.* 2006;

Kounatidis *et al.* 2008; Castrignanò *et al.* 2012; Duarte *et al.* 2015; Tamošiūnas *et al.* 2015). However, the capabilities of geospatial technologies are limited when the goal is to find relationships in ecological data, which is often highly dimensional, nonlinear and complex in nature (Cutler *et al.* 2007; Olden *et al.* 2008). Linear traditional statistical methods (e.g. generalised linear models) are not very effective in exposing the complex relationships and patterns associated with ecological datasets (De'Ath & Fabricius 2000). In contrast, ML has the ability to model non-linear relationships in ecological data without having to satisfy the restrictive assumptions associated with parametric approaches (Guisan & Zimmermann 2000; Elith *et al.* 2006; Prasad *et al.* 2006; Olden *et al.* 2008). ML, which stems from the fields of computer science and statistics (Mitchell 2006), consists of mathematical algorithms that improve through learning (Carbonell *et al.* 1983; Hamet & Tremblay 2017). Classification and regression trees (CARTs) have been widely used in ecology to model ecological relationships (De'Ath & Fabricius 2000; Thessen 2016). CARTs is a supervised learning method (Kotsiantis 2007). Supervised learning methods have numerous ecological applications in the fields of habitat and species distribution modelling, species identification, climate change and forecasting (Thessen 2016). Random forests (RF) is another supervised learning algorithm. It stems from CARTs (Breiman 2001), but has not been extensively utilised by ecologists in spite of several authors showing that it compares well with the best available classification and regression methods (Prasad *et al.* 2006; Cutler *et al.* 2007; Kampichler *et al.* 2010; Ruiz *et al.* 2010; Duque-Lazo *et al.* 2016; Mi *et al.* 2017). RF often provides high classification accuracies as it uses novel methods to determine variable importance and has the ability to model complex relationships among predictor variables (Cutler *et al.* 2007). It is also not restricted by multicollinearity in the predictor variables (Ruiz *et al.* 2010) or assumptions of the normality of the input data (Friedl & Brodley 1997), which is often present in ecological datasets. Given the ability of ML to identify complex relationships within large ecological datasets, there is a need to spatially analyse and

scrutinise regional, long-term fruit fly trap datasets such as those used by Midgarden & Lira (2006) and Enkerlin *et al.* (2016). Ruiz *et al.* (2010), for example, used RF to investigate and explain the spatial differences of the West Nile virus infection in *Culex* species mosquitoes in Illinois, USA, by considering rainfall and temperature data.

The relationships between the spatial heterogeneity of environmental factors and the spatial distribution of fruit flies at orchard or farm level have been widely studied by Israely *et al.* (2005b), Kounatidis *et al.* (2008), Nestel *et al.* (2002), Papadopoulos *et al.* (2003), Pimentel *et al.* (2014) and Puche *et al.* (2005). However, the spatial relationships between population distributions and the heterogeneity of environmental drivers (e.g. climate) at specific locations or zones on a regional scale have not been investigated. Therefore, the aim of this study was to determine the most important geographic drivers of the spatial distribution of *C. capitata* trap catches at a regional scale in a major pome fruit producing area of the Western Cape, South Africa. To this end, the spatial distribution of *C. capitata* trap catch was classified using hot spot analysis (spatial analysis), which identifies areas of relatively high (hot spots) and low trap catches (cold spots). Subjective management zones were used, and each zone was statistically compared to a range of corresponding geographic variables (possible drivers). RF was used to determine the relationships between the different variables and to classify the spatial distribution of trap catches. Finding the most important geographic divers of *C. capitata* population distributions can assist decision makers in planning and implementing better management strategies to mitigate this pest. Furthermore, to the best of my knowledge, ML and RF have not been used to investigate the effect of geographic environmental drivers on the spatial distribution of tephritid fruit flies in complex agricultural landscapes.

Background

Study area

The study area is situated in the sub-regions Elgin/Grabouw (-34.154536°, 19.029693°), Villiersdorp (-33.991976°, 19.289763°) and Vyeboom (-34.067610°, 19.109035), collectively known as EGVV. EGVV is a heterogeneous agricultural production region, but pome fruit is the predominant crop type (Figure 1).

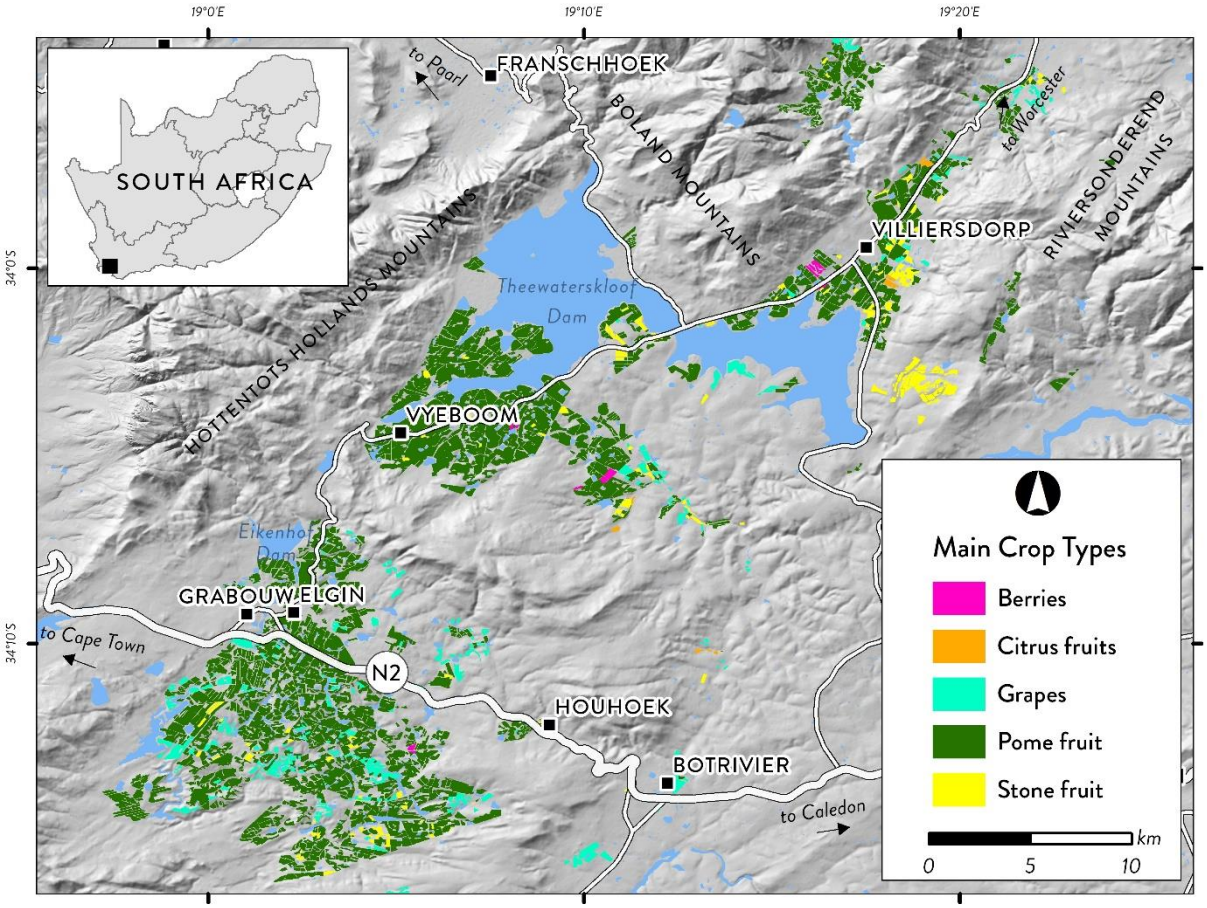

Figure 1. The spatial distribution of the main crop types planted in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa, based on the latest crop census data available.

EGVV has a Mediterranean climate with hot, dry summers and cold, wet winters, with variable temperatures and rainfall between the three sub-regions. The Elgin/Grabouw region is the coldest and wettest, with a long-term mean annual temperature of 14.9°C and a long-

term annual rainfall of 944 mm, while Villiersdorp is the hottest and driest, with a long-term mean annual temperature of 16.5°C and a long-term annual rainfall of 568 mm (Schulze 2006). Furthermore, the Hottentots Holland Mountains that border the north and western parts of Elgin/Grabouw and Vyeboom (Figure 1) are characterised by a high incidence of orographic precipitation during winter and summer months. These mountainous areas are also frequently covered in clouds, which leads to cool and damp conditions during most parts of the year (Sirgel 1985; Sieben 2003; Costandius 2005).

Ceratitis capitata monitoring data

EGVV falls under a commercial *C. capitata* AW-IPM programme operated by FruitFly Africa (Pty) Ltd. (FFA), which incorporates the Sterile Insect Technique (SIT) and Bait Application Technique (BAT). This programme furthermore operates an area-wide trapping grid for fruit flies (Barnes *et al.* 2015). During the study, all three sub-regions were part of the programme; however, Villiersdorp has left the programme in July 2016.

Fruit fly trap monitoring data for EGVV were sourced from a centralised database managed by FFA. Trap monitoring data were reported as the number of wild male, wild female and sterile male flies captured per trap per week (FTW). Trap locations in the database were geo-referenced using a hand-held GPS. All traps were monitored on a weekly to bi-weekly basis during the fruiting season (November to May), which is the period when traps in commercial fruit orchards are active (baited and monitored). Henceforth, the term season will refer to fruiting season, except when stated otherwise.

The total number of traps monitored varied among seasons and sub-regions. Using a set of selection criteria based on the consistency of monitoring effort and trap location, 399 unique traps were selected for spatial analysis. All available trap data from the 2009/2010 to the 2015/2016 seasons were initially considered, but after applying the trap selection criteria,

only data from the 2010/2011 to 2013/2014 seasons were suitable for analysis. Weekly trap data of wild *C. capitata* (male and female) were averaged per month and season (week 45–week 22), over all seasons and for individual seasons separately.

Climate and topographic data

Long-term (30 years) monthly and annual minimum and maximum temperature and rainfall data for the study area were obtained from GIS raster layers (90m resolution) developed by Van Niekerk & Joubert (2011). Other long-term climate derivatives were sourced from Schulze (2006) (Table 1).

Table 1. Summary of the climatic derivatives used in the random forest classification.

Climate derivatives	Measure	Reference
Accumulated positive chill units (May to September)	Positive chill units	(Schulze & Maharaj 2006a)
Mean number of heavy frost occurrences (annual)	Count	(Schulze & Maharaj 2006b)
Accumulated annual heat units	Degree days	(Schulze & Maharaj 2006c)
Accumulated summer heat units (October to March) (10°C)	Degree days	(Schulze & Maharaj 2006c)
Accumulated winter heat units (April to September)	Degree days	(Schulze & Maharaj 2006c)

Long-term climate data were used as it contains more climatic variability and greater climatic extremes than seasonal weather data, and therefore could account more accurately for its impacts on pest distributions (Neuvonen & Virtanen 2015). Topographic data in the form of elevation and slope gradient were derived from a high resolution (5 m) digital elevation model (DEM) (Van Niekerk 2014). Land use data, relating to urban areas, were sourced from the 2013-2014 South African Land-Cover dataset (https://egis.environment.gov.za/national_land_cover_data_sa), while crop distribution data were sourced from the Western Cape Department of Agriculture (2014).

Materials and methods

Data processing

Trap catch data from area-wide integrated pest management (AW-IPM) programmes can vary considerably from one trap to another and are thus very difficult to interpret (Miller 1956; Midgarden *et al.* 2014). Numerous factors may contribute to localized variations in trap catch, including insect activity, trap placement, trap colour, lure, weather conditions and intrinsic trap biases (Vreysen 2005). Furthermore, spatial autocorrelation between traps can also influence the interpretation of individual trap catch, because the trap catch of one trap might not be independent from the trap catch of a neighbouring trap (Sokal & Oden 1978).

Spatial autocorrelation relates to the relationship between the values of variables (e.g. trap catch) occurring in neighbouring spatial locations (Griffith 2018). When values of variables in nearby locations are similar, the spatial autocorrelation is positive, but when values of variables in nearby locations are different, negative spatial autocorrelation is experienced (Sokal & Oden 1978; Griffith 1987). Papadopoulos *et al.* (2003) investigated spatial autocorrelation of *C. capitata* trap catches in a mixed deciduous fruit orchard in northern Greece, and found that traps that were in close proximity to one another were positively correlated in terms of trap catch, while traps that were father apart from another where negatively autocorrelated.

Spatial autocorrelation and abrupt spatial variations among localized catches can be reduced by employing geostatistical techniques and by aggregation to coarser scales (Hortal *et al.* 2010). Kriging is a geostatistical technique used for predicting values at unsampled locations based on the principles of spatial autocorrelation, which measures the dependences between near and distant samples (Brenner *et al.* 1998; Childs 2004; Cox & Vreysen 2005). Ordinary kriging (OK) is a common geostatistical technique (Childs 2004) in

insect distribution investigations because it is considered to be the best linear unbiased predictor (Sciarretta & Trematerra 2014). OK was used in this study to spatially interpolate and generalize the point-based (discrete) *C. capitata* trap catch to 100 m resolution continuous raster layers. The main purpose of this process was to incorporate the relative location of each trap into the analysis (Midgarden *et al.* 2014).

Ordinary kriging was applied to weekly, monthly and seasonal (weeks 45-22) mean wild *C. capitata* catches for each season and all seasons combined. In addition, raster layers of sterile and wild males were generated for four periods including seasonal (weeks 45-22), early season (weeks 45–52), mid-season (weeks 1–11) and late season (weeks 12–22) for each season separately and all seasons combined. This temporal aggregation was done to calculate a sterile-to-wild male ratio (SWR) to investigate the contribution of sterile males released during the entire season, early season, mid-season and late season in explaining the spatial distribution of trap catches.

As suggested by Midgarden *et al.* (2014), the resulting trap catch of 100 x 100 m raster layers was further aggregated to 1 x 1 km (1 km^2) zones using Zonal Statistics in ArcGIS 10.5. A similar procedure was followed to extract the climate and topographic data. The monthly and annual long-term mean and standard deviation of all climate layers were calculated for all the raster cells within each 1 km^2 zone, while the standard deviation was calculated to measure the variation of the climatic variables within each zone. The topographic characteristics of each zone were described by calculating the mean, maximum, minimum, range and standard deviation of elevation and slope (Ruiz *et al.* 2010). Zones without trap catch records were discarded, leaving 233 zones suitable for analysis.

The SWR was calculated by dividing the mean number of sterile males by the mean number of wild males per zone. No spatial data on aerial bait application were available and were thus not included in the analysis.

To quantify and investigate the possible effects of different landscape factors on the spatial distribution of *C. capitata* trap catch, the minimum Euclidian distance (proximity measure) from urban areas and different crop types to each zone was calculated.

Spatial Analysis

Aggregated and summarised *C. capitata* trap catch for each zone were classified using the Optimised Hotspot Analysis (OHA) tool in ArcGIS 10.5 (Kounatidis *et al.* 2008). OHA identifies statistically significant spatial clustering in the data using the Global Moran's I to measure spatial autocorrelation (based on each zones' location and value). The autocorrelation is employed to determine the scale of analysis which is then used in the Getis-Ord Gi* statistic to determine the intensity of the clustering of high or low values (Getis & Ord 1992a; Ord & Getis 1995). The Gi* statistic automatically corrects for multiple testing and spatial dependence making use of the false discovery rate correction method (Caldas de Castro & Singer 2006). For a detailed description of how the OHA is conducted within the ArcGIS 10.5. Software, refer to ESRI (2017).

The OHA classified each zone as either a statistically significant hot spot or cold spot or as a non-significant zone. Statistically significant zones were binned into three groups, which reflect hot- and cold spots (HCSs) at a 99, 95 and 90 percent confidence level, while non-significant zones were also binned. To avoid class imbalances during the ML analysis HCSs at the 99, 95 and 90 percent confidence level, were binned together into a hot spot group and a cold spot group, while the non-significant zones were grouped together, resulting in three classes used in the ML analysis.

The OHA was conducted to characterise the monthly and seasonal spatial distribution of wild *C. capitata* trap catch over all seasons and for each season separately by identifying areas where high (hot spots) and low (cold spots) *C. capitata* trap catch cluster in space. This allowed for an investigation into the underlying spatial processes that may be driving the HCSs of *C. capitata* population trap catch in the study area. For visual analysis and data interpretation, HCSs maps were created. All data processing, spatial analysis and map making were conducted using ArcGIS 10.5 software.

Machine Leaning

The RF ML algorithm was used to classify the HCSs of *C. capitata* trap catch using a set of geographic variables as predictors. A summary of the geographic variables used are shown in Table 2. RF can be described as an ensemble of decision/classification trees (Breiman 2001; Liaw & Wiener 2002). RF requires two parameters, namely the number of trees (n_{tree}) to be grown in the forest and the number of random predictor variables (m_{try}) to use when splitting the data at each node (Liaw & Wiener 2002; Vermeulen & Van Niekerk 2017). RF creates n_{tree} sub-samples (bootstrap samples) from the original dataset.

Table 2. A summary of the predictor variables used in the random forest classification.

Theme	Description	Temporal scale	Number of variables
Climate	Minimum temperature	Annual and monthly	26
	Maximum temperature	Annual and monthly	26
	Mean temperature	Annual	2
	Rainfall	Annual and monthly	26
	Accumulated positive chill units	May to September	1
	Mean number of heavy frost occurrences	Annual	1
	Accumulated heat units (base 10°C)	Annual, summer (October to March) and winter (April to September)	3
Topographic	Elevation	n/a	5
	Slope	n/a	5
Landscape	Distance to urban areas	n/a	1
	Distance to grapes	n/a	1
	Distance to citrus	n/a	1
	Distance to berries	n/a	1
	Distance to stone fruit	n/a	1
Management	Sterile-to-wild ratio	Seasonal, early season, mid-season and late season	4

In our experiments RF was configured so that each bootstrap sample contains approximately 63% of the original observations, while the remaining data were used as the "out of bag" (OOB) dataset (Cutler *et al.* 2007). The algorithm grows an unpruned decision tree (classification tree), from each bootstrap sample, using only a random subset of predictor variables (m_{try}) to choose a best binary splitter of the data at each node, rather than making a selection from all the available predictor variables. Each tree grown is then used to predict the observations in the OOB dataset. RF determines overall model

accuracies by averaging the accuracies and error rates of each OOB prediction over all observations (Cutler *et al.* 2007; Rodriguez-Galiano *et al.* 2012). In addition to classification accuracies, RF, by using the Gini importance score, generates variable importance lists (VIL's) summarising the contribution of each predictor variable in conducting the classification (Breiman *et al.* 1984). The Gini importance score is an impurity measure calculated for each predictor variable used when splitting the data at each node. It measures how well a predictor variable is splitting the data at each node into the respective target classes (Breiman *et al.* 1984). The Gini importance score is an indication of how often a predictor variable was used to split the data and its overall discriminative value in classifying the data (Menze *et al.* 2009). Another method used to evaluate the importance of a variable is the permutation method, which measures the mean decrease in accuracy of the RF when the values of a predictor variable is permuted in the OOB dataset. However, it has been shown that these methods produce similar results (Strobl *et al.* 2008). Both the Gini method and the permutation method were implemented in this study to determine the importance of the predictor variables in the classification of fruit fly trap catch HCSs; however, for brevity we will only report results from the Gini method. The permutation method results for all season combined and each individual season are attached in the appendices as Table A.1 and Table A.2 respectively.

Two RF analyses were conducted. The first analysis included all available predictor variables (n = 104) to model the clustering of fruit fly trap catch, while in the second analysis, the fruit fly management predictor variables were excluded (Table 2). The OHA classifications were set as the target variable in the RF model setup. The number of trees used in the RF was set to a 1000 trees (n_{tree}) (Ruiz *et al.* 2010), while the number of predictor variables used for splitting each node (m_{try}) was set to the square root of the total number of predictor variables (Duro *et al.* 2012) (which is the default). Using a large number of trees in

RF, results in more stable classifications and variable importance measures (Cutler *et al.* 2007). RF results were formulated using SALFORD PREDICTIVE MODELER® software (www.salford-systems.com).

Statistical Analysis

The McNemar test of significance was used to compare overall model accuracies (%) of the RF models (Dietterich 1998; Foody 2004; Adedokun & Burgess 2012; Pretorius 2016). This test was configured to test statistically significant differences in overall accuracies at the 5% level of significance (Foody 2004, 2009). The statistical analyses were carried out using IBM SPSS Statistics for Windows version 25 (IBM corp 2017).

Although the spatial distribution of the monthly and seasonal *C. capitata* trap catch HCSs were analysed for each season separately, the focus of the analysis was on the combined season's dataset to better represent the long-term spatial distribution of *C. capitata* trap catch and reduce seasonal variations.

Results

Mean monthly trap catch of wild *C. capitata* varied among seasons, but generally an increase towards the end of each season was detected, with a peak from March to May. This trend was evident in each season as well as when data from all seasons were combined (Figure 2). When trap data from all seasons were combined, the variation in the mean trap catch decreased (Figure 3), which supports the decision to focus on the combined data over all seasons, rather than on the data of each season separately.

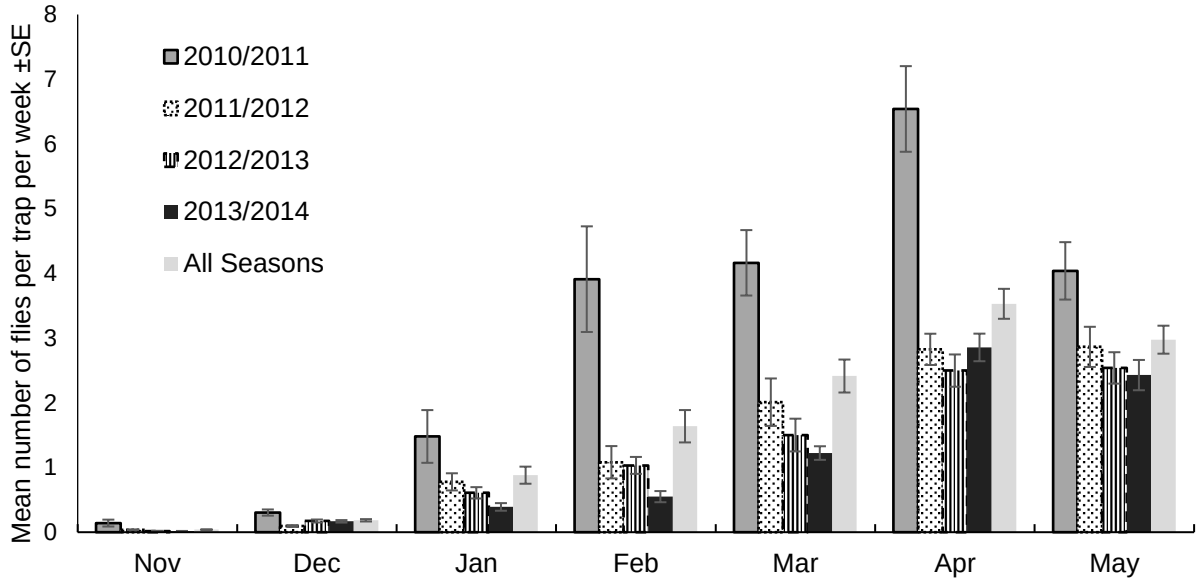

Figure 2. Mean (± SE) number of *Ceratitis capitata* per trap per week in Elgin/Grabouw, Villiersdorp, Vyeboom from November to May, for all seasons combined and each season separately in EGVV, Western Cape, South Africa.

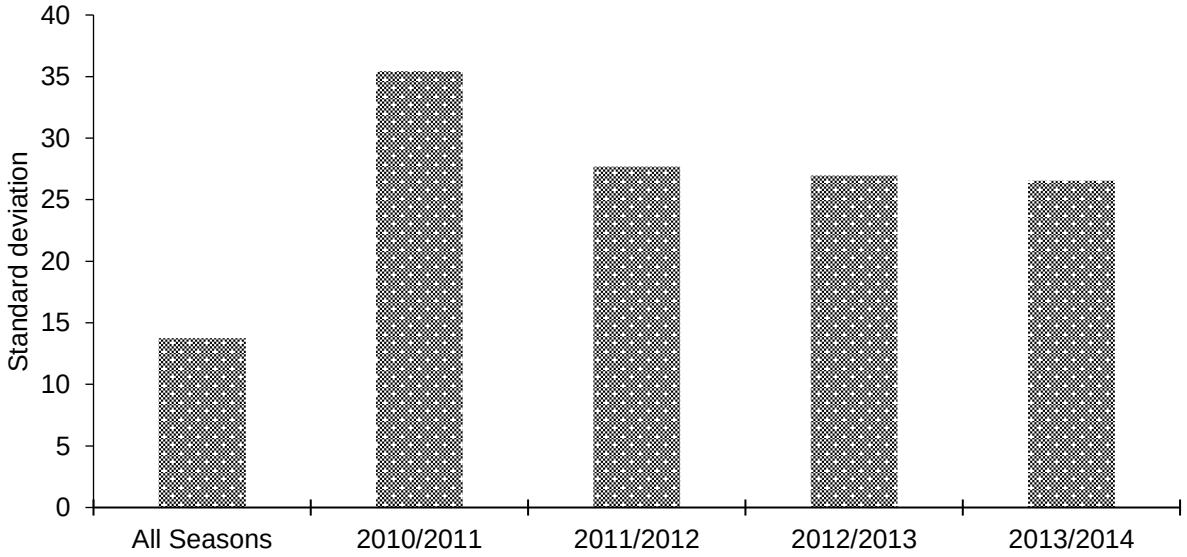

Figure 3. Standard deviation of the mean wild *Ceratitis capitata* per trap per week for all seasons combined and each season separately (2010/2011-2013/2014) in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa.

The seasonal spatial distribution of HCSs varied between seasons (Figure 4) and showed no clear pattern. However, cold spots were mainly confined to the north-westerly areas of each sub-region, while hot spots were mainly concentrated in the south-east of each sub-region (Figure 4 a-d).

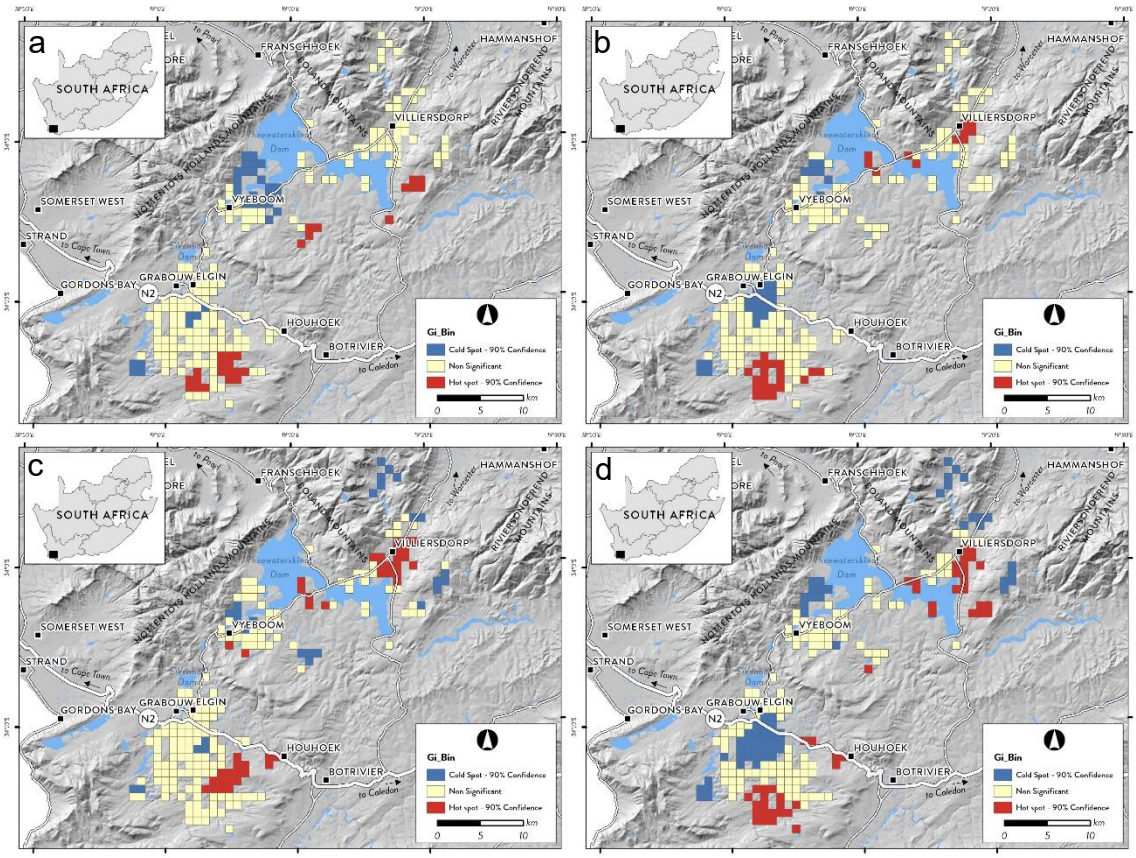

Figure 4. *Ceratitis capitata* seasonal hot- and cold spot maps for (a) 2010/2011; (b) 2011/2012; (c) 2012/2013 and (d) 2013/2014 in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa. Red zones indicate hot spots; blue zones indicate cold spots while yellow zones indicate no significant spatial pattern.

This trend was also evident in the seasonal and monthly spatial distributions of HCSs of all seasons combined (Figure 5). The extent of the HCSs varied when analysed at a monthly interval, ranging from more localized in November and February to expanding more broadly in March (Figure 5 a-g).

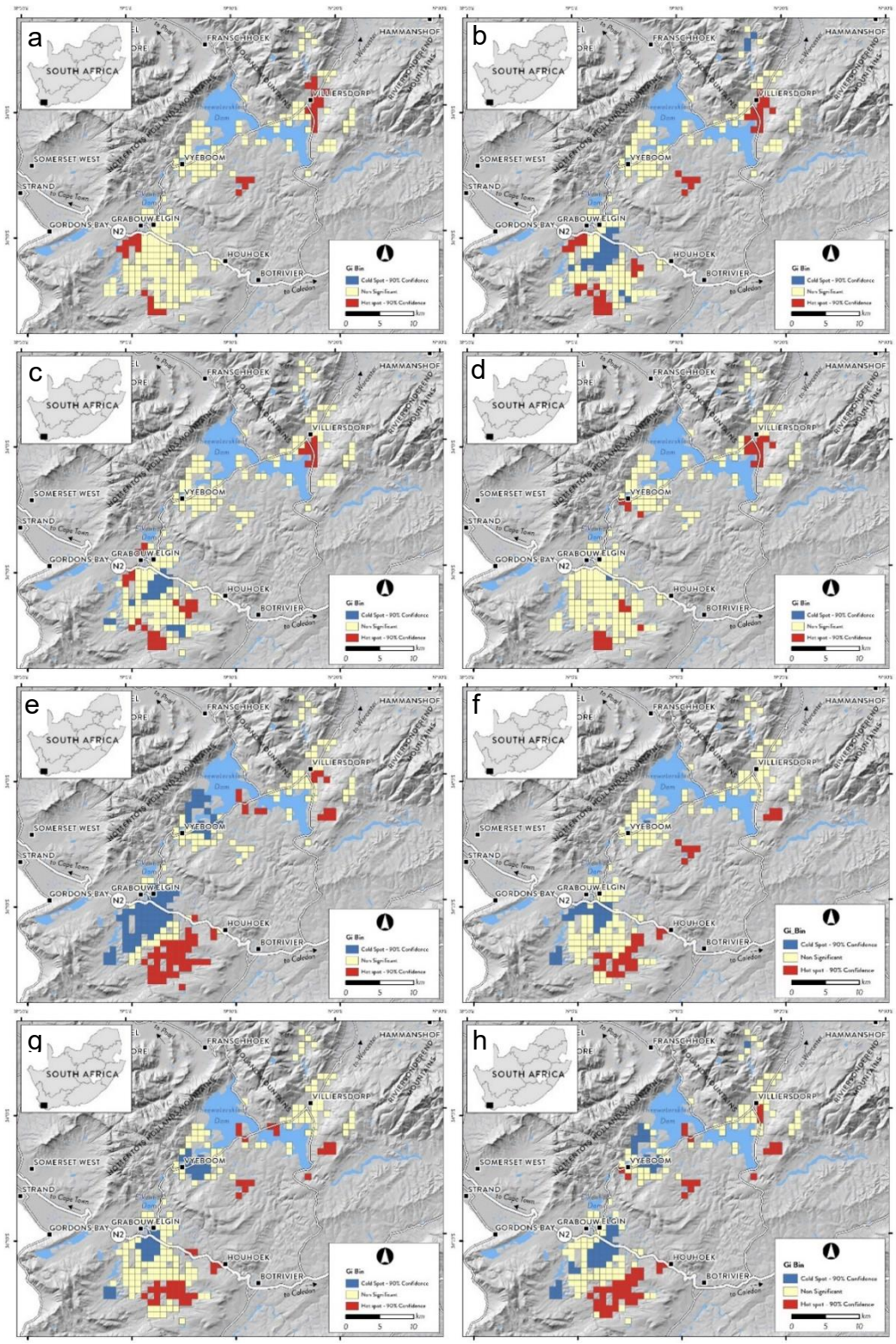

Figure 5. *Ceratitis capitata* hot- and cold spot maps for all seasons combined for (a) November; (b) December; (c) January; (d) February; (e) March; (f) April; (g) May and (h) in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa. Red zones indicate hot spots; blue zones indicate cold spots while yellow zones indicate no significant spatial pattern.

All seasons combined

Using all predictor variables as input to RF (compared to when the management variables were excluded) yielded no significant differences ($p > 0.05$) in the overall model accuracy, except for the February models where there was a significant difference ($p < 0.031$) between the overall model accuracies of the two analyses (Figure 6). In the interests of brevity, the remainder of the section will focus on the results when only the geographic predictors (i.e. excluding management variables) were used as input to RF.

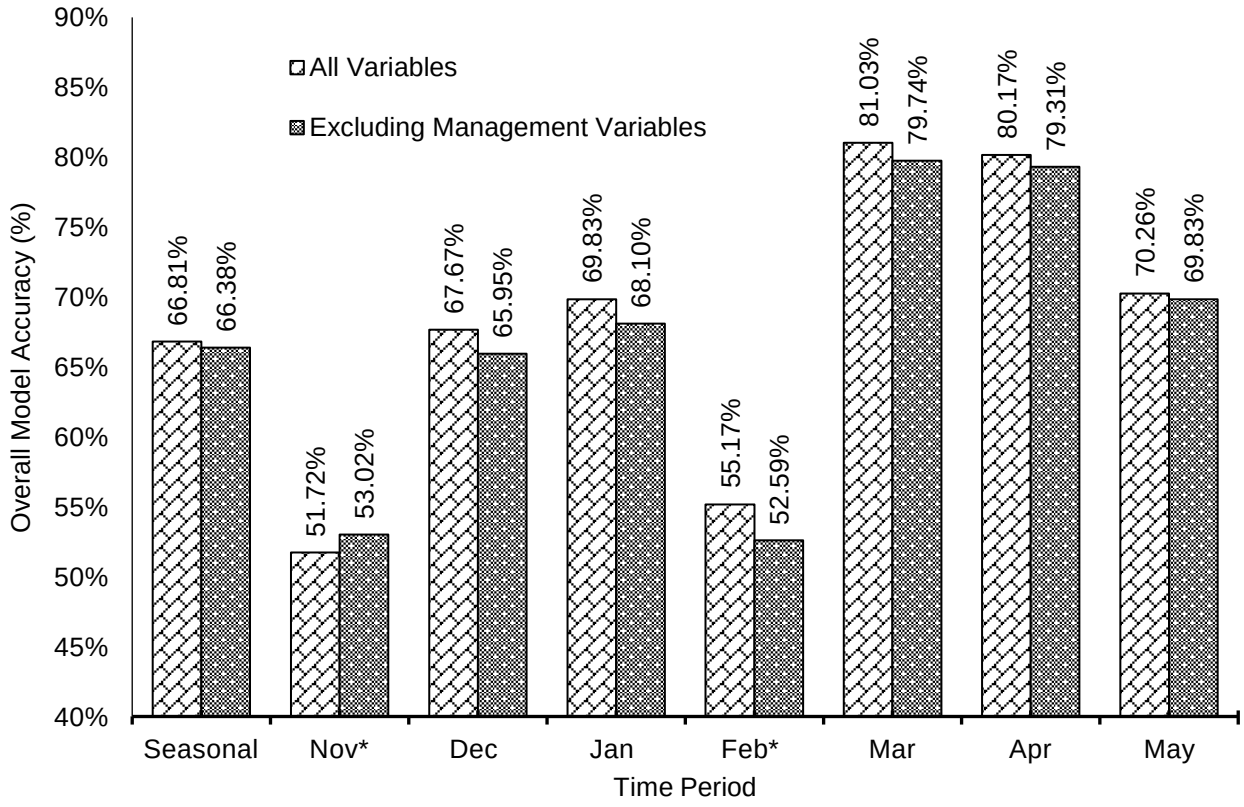

Figure 6. Overall model accuracy (%) of the random forest models when all predictor variables were used as input, compared to when management variables were excluded from the seasonal and monthly *Ceratitis capitata* trap catch models (* indicates class imbalances in the target variable).

Table 3 summarises the classification results for all tested scenarios, when data from all seasons were combined. RF produced the most accurate monthly models for March (79.74%) and April (79.31%), with December's long-term rainfall and November's long-term maximum temperature, being the most influential predictor variables respectively. In

comparison, the model accuracies of the seasonal and other monthly classifications were all in the range of 65% to 70%, with long-term rainfall being the most prominent predictor variable. The only exceptions were the November and February classifications, when the models were relatively weaker. For the November classification model, distance to urban areas emerged as the most important predictor variable, in contrast to the other models in which climatic variables were the most important.

Table 3. Summary of the random forest classification results for all seasons combined. Variable importance lists calculated using the Gini method. Bold text indicates the two most accurate models while (*) indicates class imbalances in the target variable.

Season	Time period	Overall model accuracy (%)	Most important variables in model (importance score)
All seasons combined	Seasonal	66.38%	Rainfall Dec. (100), ann. (77), May (73)
	November	53.02%*	Dist. urban (100) | Rainfall Mar. (73), Dec. (67)
	December	65.95%	Rainfall Jan. (100), Nov. (89), May (87)
	January	68.10%	Rainfall Mar. (100), Feb. (88) | Min. temp May (60)
	February	52.59%*	Rainfall Oct. (100) | Max. temp Aug. (81) | Rainfall Apr. (75)
	March	**79.74%**	**Rainfall Dec. (100), June (89), ann. (89)**
	April	**79.31%**	**Max. temp Nov. (100), ann. (89) | Rainfall Mar. (87)**
	May	69.83%	Rainfall Nov. (100), ann. (99), May (86)

Individual seasons

When the results of the individual seasons are considered (Table 4), the models of March, April and May were generally the strongest, with accuracies of about 70%. Long-term rainfall and maximum and minimum temperatures were the most influential predictors. Within all (monthly and seasonal) models of the individual seasons and the combined seasons, long-term rainfall consistently featured as the most important predictor variable, with long-term minimum and maximum temperatures also contributing to some of the models. The only exception to this pattern occurred during 2012/2013, when distance to other fruit kinds, the number of heavy frost occurrences and maximum elevation emerged as the most important predictor variables.

Table 4. Summary of the random forest classification results for individual seasons. Variable importance lists calculated using the Gini method. Bold text indicates the two most accurate models per season while (*) indicates class imbalances in the target variable.

Season	Time period	Overall model accuracy (%)	Most important variables in model (importance score)
2010/2011	Seasonal	76.72%	Rainfall ann. (100), Sept. (95), May (92)
	November	45.69%*	Rainfall Mar. (100), July (89) | Dist. urban (81)
	December	67.24%	Max. temp. Feb. (100) | Rainfall Dec. (90) | Max. temp. Dec. (80)
	January	46.98%*	Rainfall June (100), Aug. (76), Nov. (72)
	February	52.59%*	Rainfall June (100), Aug. (97), May (82)
	March	**80.17%**	**Rainfall Dec (100), Sept. (84), ann. (62)**
	April	**84.91%**	**Rainfall Oct (100), ann. (88), Sept. (84)**
	May	64.22%	Rainfall Mar. (100), Apr. (77), Oct. (75)
2011/2012	Seasonal	67.24%	Rainfall Aug. (100), June (96), Dec. (94)
	November	71.12%	Rainfall June (100), Nov. (92), Jan. (76)
	December	71.55%	Rainfall Jan. (100), Aug. (97), May. (94)
	January	71.55%	Rainfall Mar. (100), Oct. (56) | Max. temp. Jan. (55)
	February	56.90%*	Rainfall July (100), May. (75), June (70)
	March	**77.59%**	**Rainfall June (100), Aug. (87), May (84)**
	April	**78.02%**	**Rainfall Apr (100), Oct. (75) | Max. temp. June (62)**
	May	76.29%	Rainfall May. (100), Dec. (63), ann. (61)
2012/2013	Seasonal	71.12%	Max. elev. (100) | Min. elev. (89) | HeatU. Summer (81)
	November	**77.59%**	**Dist. grapes (100) | Heavy frost occ. (80) | HeatU. Summer (68)**
	December	68.53%	Dist. grapes (100) | Rainfall Jan. (96), Mar. (78)
	January	63.79%	Rainfall ann. (100), May (85), Sept. (84)
	February	**75.86%**	**Heavy frost occ. (100) | Positive chill units (91) | Rainfall ann. (77)**
	March	73.71%	Dist. citrus (100) | Rainfall Jan. (99), Mar. (80)
	April	71.12%	Max. temp. July (100), June (96) | Rainfall Jan. (85)
	May	70.26%	Rainfall Sept. (100), Dec. (80) | Max. temp. June (79)
2013/2014	Seasonal	73.71%	Rainfall May. (100), June (99), Sept. 94)
	November	62.50%	Rainfall Feb. (100) | Min. elev. (66) | Min. temp. Aug. (58)
	December	62.93%	Rainfall Jan. (100), July (70), Feb. (60)
	January	75.00%	Elev. (100) | Min. elev. (84) | Max. temp. June (64)
	February	37.50%*	Positive chill units (100) | HeatU. Winter (70) | Rainfall ann. (66)
	March	68.10%	Rainfall June (100), Aug. (76), May (73)
	April	**75.43%**	**Rainfall Jan. (100), Sept. (74), Dec. (74)**
	May	**75.43%**	**Rainfall Dec. (100), June (83), Sept. (76)**

Given the overwhelming importance of long-term mean temperature (°C) as well as the long-term mean annual rainfall (mm) in the models, maps of these variables (Figure 7) were produced to better understand the results. The maps clearly show that there were important climatic differences between zones in the north-west and zones in the south-east, which resonates well with the patterns observed in Figure 5.

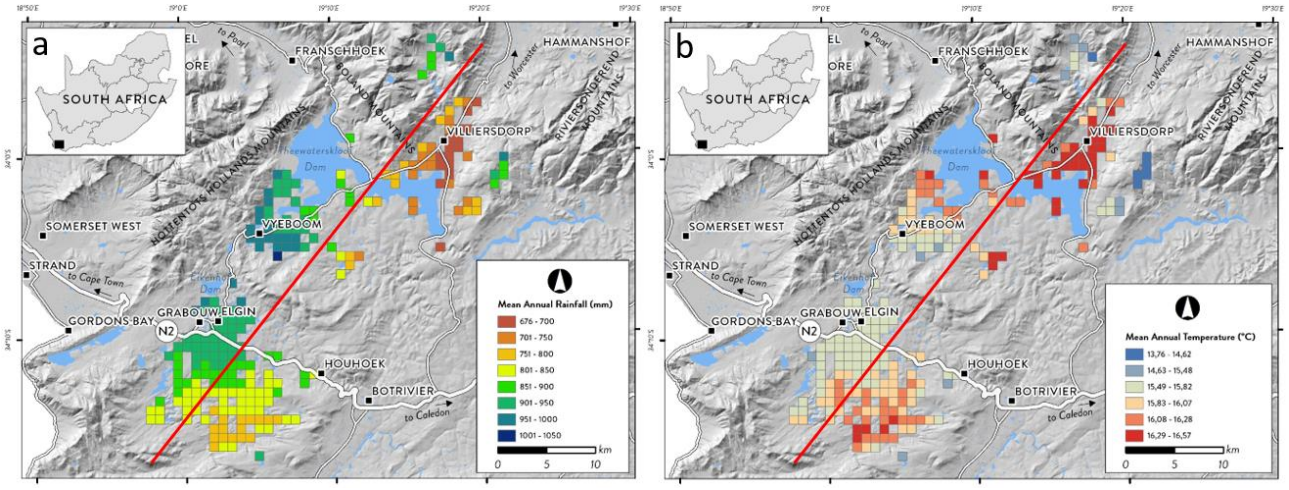

Figure 7. Mean annual long-term rainfall (a) and temperature (b) per zone in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa. The red line in the maps serve to point out the clear division between zone values in the north-west and the south-east.

The fact that long-term rainfall predominantly emerged as the most important geographic driver of *C. capitata* HCSs in most models, but not in the 2012/2013 season models, compelled us to investigate the deviation of the seasonal rainfall from the long-term rainfall. Seasonal rainfall data was obtained from two automatic weather stations in the study area and were compared to the long-term weather data extracted at the coordinates of each weather station. Results represented in Figure 7a, indicated that the long-term rainfall pattern was likely a good representation of the rainfall conditions during 2010 to 2014, as the deviation from the long-term rainfall was generally small as observed in Figure 8. The only exception was from August to November in 2012 and 2013, when the monthly rainfall deviated strongly from the mean long-term rainfall.

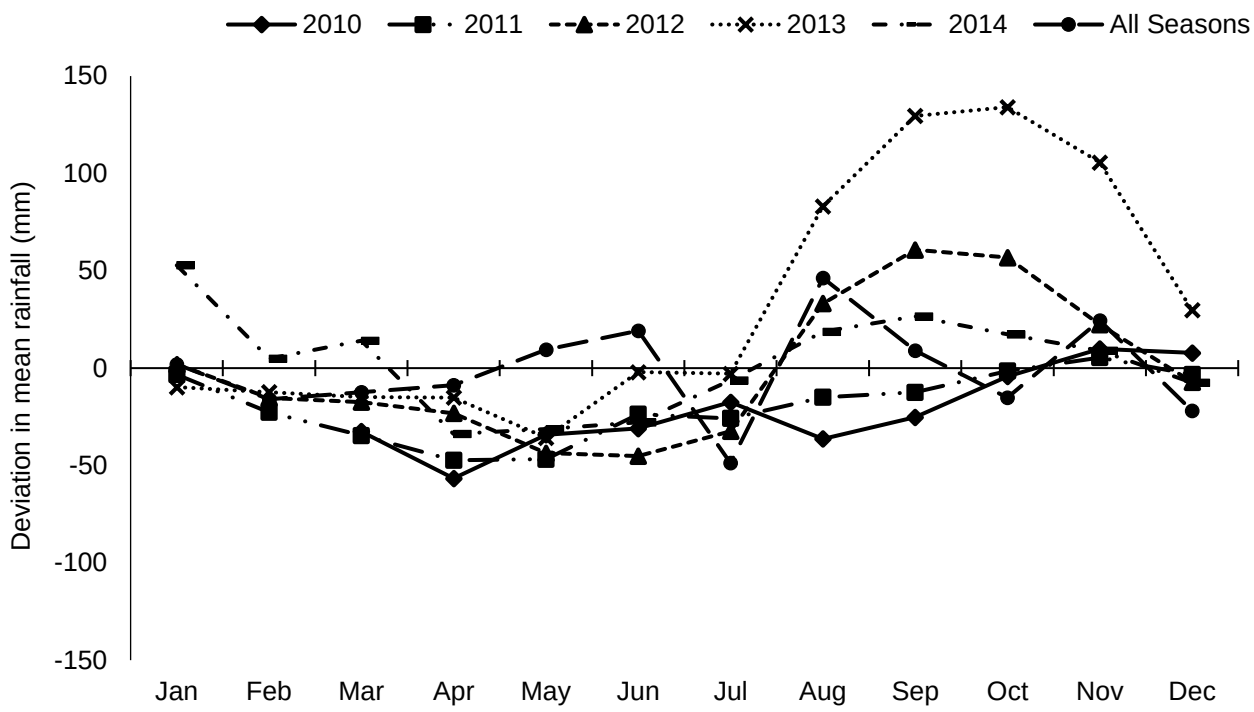

Figure 8. The three-month moving average of the monthly rainfall (mm) per year, indicating the deviation from the long-term mean rainfall in Elgin/Grabouw, Villiersdorp, Vyeboom, Western Cape, South Africa

Discussion

The concentration of hot spots in the south-eastern parts and cold spots in the north-western parts of each sub-region of EGVV can likely be ascribed to spatial heterogeneity of the climate, especially rainfall, and to a lesser degree temperature, within these areas. It is clear that the south-eastern parts of each sub-region are generally much drier but also slightly warmer compared to their north-western counterparts, confirming the hypothesis of the previous chapter. This geographical division of HCSs might be due to *C. capitata* populations being physiologically better adapted to drier and hotter climates (Duyck *et al.* 2006; Nyamukondiwa & Terblanche 2010). *Ceratitis rosa* Karsh (Diptera: Tephritidae), subsequently split into two species and renamed *Ceratitis quilicii* De Meyer, Mwatawala and Virgilio (Diptera: Tephritidae) (De Meyer *et al.* 2016), which is assumed to be present in the Western Cape Province, also occurs within the study area (Manrakhan & Addison 2014). This species is known to prefer wetter and cooler climates (Duyck *et al.* 2006; DeVilliers *et*

al. 2013) and it has been shown that *C. capitata* and *C. quilicii* segregate ecologically and geographically (Duyck *et al.* 2006), which was ascribed to niche-dependent competition, but also the physiological response of each species to climate. Unfortunately, no area-wide, long-term trap data were available for *C. quilicii* in EGVV. Furthermore, high rainfall has also been shown to contribute to adult mortality of *C. capitata* (Peñarrubia-María *et al.* 2012b) and, mortality of the pupal stage when temporarily submerged in water (Duyck *et al.* 2006b). *Ceratitis capitata* populations in the higher rainfall areas in EGVV might consequently have been subjected to this mortality, which could have formed the basis of the spatial distribution of HCSs of the trap catches.

The spatial distribution of *C. capitata* populations has been shown to be related to host availability and suitability (Nestel *et al.* 2002; Sciarretta & Trematerra 2011). It is likely that these factors contributed to the variation in the spatial distribution of the HCSs within the limits of the south-east/north-west split during different times (months), given that EGVV is a heterogeneous fruit production region, with different crop types and cultivars ripening at different periods. The HCSs were most prominent in March, which coincides with the time when the majority of hosts fruits are ripening in the commercial orchards in EGVV (Manrakhan & Addison 2014). With the absence of hosts in the commercial orchards, flies disperse in search of alternate hosts (Myburgh 1956; Barnes 2008). A reduction in the number of HCSs could be seen in April and May.

The fact that overall model accuracy of the RF models did not change significantly when the management variables were excluded suggests that SWR was not a significant driver of the long-term seasonal or monthly spatial distribution of *C. capitata* HCSs in EGVV. This brings the effectiveness of the application process of the sterile flies into question and suggests that the quantum, spatial distribution and method of release may have to be revisited. The

fitness of the sterile flies and how well they compete with wild males should also be considered. Biotic processes such as competition between individuals for mates and other resources can play an important role in determining the spatial distribution of a species at a fine spatial scale, but as the scale becomes coarser, the effect on the spatial distribution becomes negligible (Hortal *et al.* 2010; Soberón 2010). The functioning of SIT is largely based on the competition between sterile and wild male flies (McInnis *et al.* 1994); therefore, the effect of the SWR may have been concealed by the relatively coarse scale at which this study was conducted. The significant difference in the model accuracies in the February models indicates that SWR could drive the spatial distribution of HCSs; however, on a regional scale in EGVV the effect was minor.

The RF results indicated that the HCSs of March and April, when data from all seasons were combined related strongly to the climate of preceding months. Peak levels in *C. capitata* trap catch, indicating higher population levels, have been reported to be higher when the preceding winters were milder with less rainfall (Peñarrubia-María *et al.* 2012). This makes sense when considering that low winter temperatures and high rainfall cause adult and pupal mortality (Papadopoulos *et al.* 1996; Duyck *et al.* 2006; Peñarrubia-María *et al.* 2012). High winter mortality would have resulted in smaller populations in spring and early summer, thus leading to lower peak population levels later during the season (Escudero-Colomar *et al.* 2008; Enkerlin *et al.* 2016), impacting on the spatial distribution of the HCSs, which was based on trap catch.

The fact that long-term climate variables, in particular rainfall, were the most important in all the RF models, suggests that the long-term climate data represented rainfall of the individual years (2010-2014) well. The only exception was the 2012/2013 season, in which long-term rainfall was not the main driver. Recorded rainfall during the spring months (August–

November) of 2012, which preceded the 2012/2013 season, deviated substantially from the long-term averages. This could have resulted in the long-term spring rainfall not being a good predictor of the 2012/2013 season's HCSs. However, this theory does not hold for the 2013/2014 season when the recorded rainfall of the 2013 spring months deviated even more dramatically from the long-term averages but were still the most important predictor variable. The available climate data consequently do not conclusively explain the 2012/2013 season deviation in the RF results.

Conclusion

HCSs based on fruit fly trap catch, were successfully classified during different seasonal periods using OHA. Within a very limited region, a clear split in the long-term spatial distribution of HCSs was identified, indicating underlying spatial processes at work. The variability of monthly and seasonal HCSs highlighted the need for longer-term trapping data on a regional scale to sufficiently quantify these underlying spatial processes. RF identified long-term rainfall as the most prominent predictor variable in almost all the RF models. Furthermore, with the use of long-term environmental data, RF could successfully classify HCSs with up to an 80% accuracy. This highlights the value of long-term environmental data representing seasonal variations for improving our understanding of the relationships between the spatial distribution of these pests and their environment. The information gained will contribute to a better understanding of the role that geographic variables play in the spatial distribution of *C. capitata* trap catch. This information is invaluable in area-wide fruit fly management programmes, especially those that incorporate SIT, as it will allow for more precise spatial planning, which could lead to better programme performance and reduced costs. It also presents a more proactive approach in assessing risk in terms of stable geographic characteristics of an area, rather than focusing on highly variable seasonal factors. AW-IPM programme managers are encouraged to scale-up their management

actions in specific areas (hot spots) while still continuing with the current blanket approach. Future research should test whether the same geographic drivers of *C. capitata* HCSs emerge as important in other fruit producing regions of South Africa, and monitoring should expand to include data for *C. quilicii*, which is currently not available The geospatial approach used in this study may provide a good foundation for such work.

References

ABD EL WAHAB, M.K., EL SHAZLY, M.A., EL DEEB, M.A. & GHIYATY, M.F. 2006. Spatio-temporal monitoring of Mediterranean Fruit Fly *Ceratitis capitata* presence by Precision Agriculture Techniques. *Biological Engineering*: 1178–1186.

ADEDOKUN, O.A. & BURGESS, W.D. 2012. Analysis of Paired Dichotomous Data: A Gentle Introduction to the McNemar Test in SPSS. *Journal of MultiDisciplinary Evaluation* **8**(17): 125–131.

BARNES, B. 2008. Why have fruit flies become so troublesome? *South African Fruit Journal* **7**(4): 56–57.

BARNES, B.N., HOFMEYR, J.H., GROENEWALD, S., CONLONG, D.E. & WOHLFARTER, M. 2015. The sterile insect technique in agricultural crops in South Africa : a metamorphosis .... but will it fly ? *African Entomology* **23**(1): 1–18.

BREIMAN, L. 2001. Random forests. *Machine Learning* **45**(1): 5–32.

BREIMAN, L., FRIEDMAN, J., OLSHEN, R. & STONE, C. 1984. *Classification and Regression Trees*. Wadsworth, Belmont, California.

BRENNER, R.J., FOCKS, D.A., ARBOGAST, R.T., WEAVER, D.K. & SHUMAN, D. 1998. Practical use of Spatial Analysis in Precision Targeting for Integrated Pest Management. *American Entomologist* **44**(2): 79–102.

CALDAS DE CASTRO, M. & SINGER, B.H. 2006. Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Test in Local Statistics of Spatial Association. *Geographical Analysis* **38**: 180–208.

CARBONELL, J.G., MICHALSKI, R.S. & MITCHELL, T.M. 1983. An overview of machine learning. In: Michalski, R.S., Carbonell, J.G. & Mitchell, T.M. (Eds.) *Machine learning: An Artificial Intelligence Approach*. 3–23. Springer, Palo Alto, California.

CASTRIGNANÒ, A., BOCCACCIO, L., COHEN, Y., NESTEL, D., KOUNATIDIS, I., PAPADOPOULOS, N.T., DE BENEDETTO, D. & MAVRAGANI-TSIPIDOU, P. 2012. Spatio-temporal population dynamics and area-wide delineation of *Bactrocera oleae* monitoring zones using multi-variate geostatistics. *Precision Agriculture* **13**(4): 421–441.

CHILDS, C. 2004. Interpolating surfaces in ArcGIS spatial analyst. *ArcUser, July-September* **3235**: 569.

COSTANDIUS, E. 2005. The Landdroskop area in the Hottentots Holland Mountains as a refugium for melanistic lizard species : an analysis for. Stellenbosch University.

COX, J.S.H. & VREYSEN, M.J.B. 2005. Use of Geographic Information Systems and Spatial Analysis in Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 453–477. Springer, Dordrecht.

CUTLER, D.R., EDWARDS, T.C., BEARD, K.H., CULTLER, A., HESS, K.T., GIBSON, J. & LAWLER, J.J. 2007. Random Forests for Classification in Ecology. *Ecological Society of America* **88**(11): 2783–2792.

DE'ATH, G. & FABRICIUS, K.E. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* **81**(11): 3178–3192.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DIETTERICH, T.G. 1998. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation* **10**(7): 1895–1923.

DUARTE, F., CALVO, M.V., BORGES, A. & SCATONI, I.B. 2015. Geostatistics applied to the study of the spatial distribution of insects and its use in integrated pest management. *Revista agronómica del noroeste argentino* **35**(2): 9–20.

DUQUE-LAZO, J., VAN GILS, H., GROEN, T.A. & NAVARRO-CERRILLO, R.M. 2016. Transferability of species distribution models : The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. *Ecological Modelling* **320**: 62–70. Elsevier B.V.

DURO, D.C., FRANKLIN, S.E. & DUBÉ, M.G. 2012. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *International Journal of Remote Sensing* **33**(14): 4502–4526.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

ELITH, J., GRAHAM, C.H., ANDERSON, R.P., DUDÍK, M., GUISAN, A., HIJMANS, R.J., HUETTMANN, F., LEATHWICK, J.R., LEHMANN, A., LI, J., LOHMANN, L.G., LOISELLE, B.A., MANION, G., NAKAMURA, M., NAKAZAWA, Y., OVERTON, J.M., TOWNSEND, A., PHILLIPS, S.J., RICHARDSON, K., SCACHETTI-PEREIRA, R., ROBERT, E., SOBERÓN, J. & WILLIAMS, S. 2006. Novel Methods Improve Prediction of Species' Distributions from Occurrence Data. *Ecography* **29**(2): 129–151.

ENKERLIN, W., VILLASEÑOR, A., FLORES, S., MIDGARDEN, D., LIRA, E., RENDON, P., HURLEY, J., SALAZAR, E., MENDEZ, W., CASTAÑEDA, R., COTOC, E., ZAVALA, J.L., CELEDONIO, H. & GUTIERREZ-RUELAS, J.M. 2016. Descriptive analysis of the factors affecting the population dynamics of the Mediterranean fruit fly (*Ceratitis capitata*, Wied.) in coffee areas located in Guatemala and its implications in IPM Strategies. In: Sabater-Muñoz, B., Vera, M.T., Pereira, R. & Orankanok, W. (Eds.) *Proceedings IX International Symposium on Fruit Flies of Economic Importance, 12-14 May 2014*. 46–63.

ESCUDERO-COLOMAR, L.A., VILAJELIU, M. & BATLLORI, L. 2008. Seasonality in the occurrence of the Mediterranean fruit fly [*Ceratitis capitata* (Wied.)] in the north-east of Spain. *Journal of Applied Entomology* **132**(9–10): 714–721.

ESRI. 2017. How Optimised Hot Spot Analysis Works.

FOODY, G.M. 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering & Remote Sensing* **70**(5): 627–633.

FOODY, G.M. 2009. Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing* **30**(20): 5273–5291.

FRIEDL, M. A. M.A. & BRODLEY, C.E.C.E. 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment* **61**(3): 399–409.

GETIS, A. & ORD, J.K. 1992. The Analysis of Spatial Association. *Geographical Analysis* **24**(3): 189–206.

GRIFFITH, D.A. 1987. Spatial autocorrelation. *A Primer. Washington DC: Association of American Geographers*.

GRIFFITH, D.A. 2018. What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *L'Espace geographique* **21**(3): 265–280.

GUISAN, A. & ZIMMERMANN, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.

GUTIERREZ, A.P. & PONTI, L. 2011. Assessing the invasive potential of the Mediterranean fruit fly in California and Italy. *Biological Invasions* **13**(12): 2661–2676.

HAMET, P. & TREMBLAY, J. 2017. Artificial intelligence in medicine. *Metabolism: Clinical and Experimental* **69**: S36–S40. Elsevier Inc.

HENDRICHS, J. & HENDRICHS, M.A. 1990. Mediterranean Fruit Fly (Diptera: Tephritidae) in Nature: Location and Diel Pattern of Feeding and Other Activities on Fruiting and Nonfruiting Hosts and Nonhosts. *Annals of the Entomological Society of America* **83**(3): 632–641.

HENDRICHS, J., KENMORE, P. & ROBINSON, A.S. 2007. Area-wide integrated pest management (AW-IPM): principles, practice and prospects. In: Vreysen, M.J.B., Robinson, A.S. & Hendrichs, J. (Eds.) *Area-wide control insect pests: from research to field implementation*. 3–33. Springer, Dordrecht.

HORTAL, J., ROURA-PASCUAL, N., SANDERS, N.J. & RAHBEK, C. 2010. Understanding (insect) species distributions across spatial scales. *Ecography* **33**(1): 51–53.

IBM CORP. 2017. IBM SPSS Statistics for Windows, Version 25. Armonk, NY: IBM Corp.

ISRAELY, N., ZIV, Y. & GALUN, R. 2005a. Metapopulation Spatial – Temporal Distribution Patterns of Mediterranean Fruit Fly (Diptera: Tephritidae) in a Patchy Environment. *Annals of the Entomological Society of America* **98**(3): 302–308.

ISRAELY, N., ZIV, Y. & OMAN, S.D. 2005b. Spatiotemporal Distribution Patterns of Mediterranean Fruit Fly (Diptera : Tephritidae) in the Central Region of Israel. *Annals of the Entomological Society of America* **98**(1): 77–84.

KAMPICHLER, C., WIELAND, R., CALMÉ, S., WEISSENBERGER, H. & ARRIAGA-WEISS, S. 2010. Classification in conservation biology: A comparison of five machine-learning methods. *Ecological Informatics* **5**(6): 441–450.

KOTSIANTIS, S.B. 2007. Supervised Machine Learning: A Review of Classification Techniques. *Informatica* **31**: 249–268.

KOUNATIDIS, I., PAPADOPOULOS, N.T., MAVRAGANI-TSIPIDOU, P., COHEN, Y., TERTIVANIDIS, K., NOMIKOU, M. & NESTEL, D. 2008. Effect of elevation on spatio-temporal patterns of olive fly (*Bactrocera oleae*) populations in northern Greece. *Journal of Applied Entomology* **132**: 722–733.

LIAW, A. & WIENER, M. 2002. Classification and Regression by randomForest. *R news* **2/3**(December): 18–22.

LINDSAY, K.R. 2010. The impacts of climate change on the summerfruit industry with respect to insect pest incursions. Lincoln University.

MANRAKHAN, A. & ADDISON, P. 2014. Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science* **70**(4): 651–660.

MAZZI, D. & DORN, S. 2012. Movement of insect pests in agricultural landscapes. *Annals of Applied Biology* **160**(2): 97–113.

MCINNIS, D.O., TAM, S., GRACE, C. & MIYASHITA, D. 1994. Population Suppression and Sterility Rates Induced by Variable Sex Ratio, Sterile Insect Releases of *Ceratitis capitata* (Diptera: Tephritidae) in Hawaii. *Annals of the Entomological Society of America* **87**(2): 231–240.

MENZE, B.H., KELM, B.M., MASUCH, R., HIMMELREICH, U., BACHERT, P., PETRICH, W. & HAMPRECHT, F.A. 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* **10**(213): 1–16.

DE MEYER, M., MWATAWALA, M., COPELAND, R.S. & VIRGILIO, M. 2016. Description of new *Ceratitis* species (Diptera: Tephritidae) from Africa, or how morphological and DNA data are complementary in discovering unknown species and matching sexes. *European Journal of Taxonomy* **233**: 1–23.

MI, C., HUETTMANN, F., GUO, Y., HAN, X. & WEN, L. 2017. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ* **5**: e2849.

MIDGARDEN, D. & LIRA, E. 2006. Ecological relationship of Medfly and coffee in guatemala and Mexico. *Fruit Flies of Economic Importance: From Basic to Applied Knowledge Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*: 241–247.

MIDGARDEN, D.G., LIRA, E. & SILVER, M. 2014. Spatial Analysis of Tephritid Fruit Fly Traps. In: Shelly, T., Epsky, N., Jang, E.B., Reyes-Floris, J. & Vargas, R. (Eds.) *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. 277–320. Springer, Dordrecht.

MILLER, G.A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Phychological Review* **63**(2): 81–96.

MITCHELL, T.M. 2006. The discipline of machine learning. Carnegie Mellon University, School of Computer Science, Machine Learning Department.

MYBURGH, A.C. 1956. Bionomics and control of the fruit flies, *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.), in the Western Cape. University of Stellenbosch.

NESTEL, D., KATSOYANNOS, B.I., NEMNY-LAVY, E., MENDEL, Z. & PAPADOPOULOS, N.T. 2002. Spatial analysis of Medfly populations in heterogeneous landscapes. *Proceedings of 6th International Fruit Fly Symposium*. 35–43. Stellenbosch, South Africa.

NEUVONEN, S. & VIRTANEN, T. 2015. Abiotic factors, climatic variability and forest insect pests. In: Björkman, C. & Niemelä, P. (Eds.) *Climate change and insect pests*. 154–172. CABI Publishing.

VAN NIEKERK, A. 2014. Stellenbosch University Digital Elevation Model (SUDEM), 2013 Edition (v13.24). Stellenbosch, South Africa.

VAN NIEKERK, A. & JOUBERT, S.J. 2011. Input variable selection for interpolating high-resolution climate surfaces for the Western Cape. *Water SA* **37**(3): 271–280.

NYAMUKONDIWA, C. & TERBLANCHE, J.S. 2010. Within-generation variation of critical thermal limits in adult Mediterranean and Natal fruit flies *Ceratitis capitata* and *Ceratitis rosa*: Thermal history affects short-term responses to temperature. *Physiological Entomology* **35**(3): 255–264.

OLDEN, J.D., LAWLER, J.J. & POFF, N.L. 2008. Machine Learning Methods Without Tears: A Primer for Ecologists. *The Quarterly Review of Biology* **83**(2): 171–193.

ORD, J.K. & GETIS, A. 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* **27**(4): 286–306.

PAPADOPOULOS, N.T., CAREY, J.R., KATSOYANNOS, B.I. & KOULOUSSIS, N.A. 1996. Overwintering of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Northern Greece. *Annals of the Entomological Society of America* **48**(4): 526–534.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PEÑARRUBIA-MARÍA, E., AVILLA, J., ESCUDERO-COLOMAR, L.A., PENAEEUBIA-MARIA, E., AVILLA, J. & ESCUDERO-COLOMAR, L.A. 2012. Survival of wild adults of *Ceratitis capitata* (Wiedemann) under natural winter conditions in North East Spain. *Psyche* **2012**: 1–6.

PIMENTEL, R., LOPES, D.J.H., MEXIA, A.M.M. & MUMFORD, J.D. 2014. Spatial regression analysis of *Ceratitis capitata* Wiedemann (Diptera: Tephritidae) on Terceira Island, Azores. *International Journal of Pest Management* **60**(3): 217–223.

PRASAD, A.M., IVERSON, L.R. & LIAW, A. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **9**(2): 181–199.

PRETORIUS, A. 2016. Advances in Random Forests with Application to Classification. Stellenbosch University.

PUCHE, H., MIDGARDEN, D.G., OVALLE, O., KENDRA, P.E., EPSKY, N.D., RENDON, P. & HEATH, R.R. 2005. Effect of Elevation and Host Availability on Distribution of Sterile and Wild Mediterranean Fruit Flies (Diptera: Tephritidae). *Florida Entomologist* **88**: 83–90.

RODRIGUEZ-GALIANO, V.F., CHICA-OLMO, M., ABARCA-HERNANDEZ, F., ATKINSON, P.M. & JEGANATHAN, C. 2012. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment* **121**: 93–107. Elsevier Inc.

RUIZ, M.O., CHAVES, L.F., HAMER, G.L., SUN, T., BROWN, W.M., WALKER, E.D., HARAMIS, L., GOLDBERG, T.L. & KITRON, U.D. 2010. Local impact of temperature and precipitation on West Nile virus infection in Culex species mosquitoes in northeast Illinois, USA. *Parasites & Vectors* **3**(1): 19.

SCHULZE, R. 2006. *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06.

SCHULZE, R. & MAHARAJ, M. 2006a. Positive Chill Units. In: Schulze, R. (Ed.) *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06, Section 10.

SCHULZE, R. & MAHARAJ, M. 2006b. Mean Number of Occurances of Heavy Frost. In: Schulze, R. (Ed.) *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06, Section 9.

SCHULZE, R. & MAHARAJ, M. 2006c. Heat Units. In: Schulze, R. (Ed.) *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06, Section 8.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

SCIARRETTA, A. & TREMATERRA, P. 2014. Geostatistical Tools for the Study of Insect Spatial Distribution: Practical Implications in the Integrated Management of Orchard and Vineyard Pests. *Plant Protection Science* **50**(2): 97–110.

SIEBEN, E.J.J. 2003. The riparian vegetation of the Hottentots Holland Mountains, SW Cape. Stellenbsoch University.

SIRGEL, W.F. 1985. A new subfamily of Arionidae (Mollusca, Pulmonata). *Annals of the Natal Museum* **26**(2): 471–487.

SOBERÓN, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* **10**(12): 1115–1123.

SOBERÓN, J. & NAKAMURA, M. 2009. Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences* **106**(Supplement 2): 19644–19650.

SOBERÓN, J.M. 2010. Niche and area of distribution modeling: A population ecology perspective. *Ecography* **33**(1): 159–167.

SOKAL, R.R. & ODEN, N.L. 1978. Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean society* **10**: 199–228.

STROBL, C., BOULESTEIX, A., KNEIB, T., AUGUSTIN, T. & ZEILEIS, A. 2008. Conditional variable importance for random forests. *BMC* **9**(307): 1–11.

TAMOŠIŪNAS, R., VALIUŠKAITĖ, A., JUKNA, L., TAMOŠIŪNAS, K. & ŽIOGAS, A.F. 2015. Spatial distribution patterns of apple sawfly populations in two differently managed commercial apple orchards. *Zemdirbyste-Agriculture* **102**(1): 73–80.

THESSEN, A. 2016. Adoption of Machine Learning Techniques in Ecology and Earth Science. *One Ecosystem* **1**: e8621.

VARGAS, R.I., HARRIS, E.J. & NISHIDA, T. 1983. Distribution and Seasonal Occurrence of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) on the Island of Kauai in the Hawaiian Islands. *Environmental Entomology* **12**(2): 303–310.

VERMEULEN, D. & VAN NIEKERK, A. 2017. Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates. *Geoderma* **299**: 1–12. Elsevier B.V.

VREYSEN, M.J.B. 2005. Monitoring Sterile and Wild Insects in Area-wide Integrated Pest Management Programmes. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 325–361. Springer Netherlands, Dordrecht.

WAGNER, H.H. & FORTIN, M.-J. 2005. Spatial Analysis of Landscapes : Concepts and Statistics. *Ecology* **86**(8): 1975–1987.

WESTERN CAPE DEPARTMENT OF AGRICULTURE. 2014. The Mapping of Agricultural Commodity Production and Agricultural Infrastructure in the Western Cape.

WHITE, I.M. & ELSON-HARRIS, M.M. 1992. *Fruit flies of economic significance: their identification and bionomics*. CAB International/Australian Centre for International Agricultural Research, Wallingford, UK/Canberra.

WHITTAKER, R.J., WILLIS, K.J. & FIELD, R. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography* **28**(4): 453–470.

Appendix

Table A.1 Summary of the random forest classification results for all seasons combined. Variable importance lists calculated using the permutation method. Bold text indicates the two most accurate models while (*) indicates class imbalances in the target variable.

Season	Time period	Overall model accuracy (%)	Most important variables in model (importance score)
All seasons combined	Seasonal	66.38%	Rainfall Dec. (100), May (96), Sept. (91)
	November	53.02%*	Dist. urban (100) | Rainfall Dec. (89), July (79)
	December	65.95%	Rainfall June (100), July (80), May (77)
	January	68.10%	Rainfall Mar. (100), Feb. (85) | Min. temp. June (69)
	February	52.59%*	Max. temp. ann. (100) | Max. temp. Nov. (98) | Rainfall Apr. (91)
	March	**79.74%**	**Rainfall May (100), June (89), Sept. (78)**
	April	**79.31%**	**Rainfall Sept. (100), Dec. (71) | Rainfall Feb. (70)**
	May	69.83%	Rainfall Feb. (100), Apr. (99), May (82)

Table A.2 Summary of the random forest classification results for individual seasons. Variable importance lists calculated using the permutation method. Bold text indicates the two most accurate models per season while (*) indicates class imbalances in the target variable.

Season	Time Period	Overall Model Accuracy (%)	Most important variables in model (importance score)
2010/2011	Seasonal	76.72%	Rainfall Dec. (100), ann. (85), Sept. (79)
	November	45.98%*	Rainfall June (100), Dec. (92) | Min. temp. June (92)
	December	66.81%	Rainfall Nov. (100) | Rainfall Jan. (73) | Max. temp. Feb (66)
	January	43.53%*	Rainfall June (100), July (73), Nov. (68)
	February	50.86%*	Rainfall June (100), Aug. (79), May (50)
	March	**80.17%**	**Rainfall Dec. (100), Sept. (74), June (73)**
	April	**84.05%**	**Min. temp. July (100) | Rainfall June (98), May (81)**
	May	65.52%	Min. elev. (100) | Rainfall May (98), Feb. (97)
2011/2012	Seasonal	65.95%	Rainfall Aug. (100), Dec. (99) | Min. temp. May (98)
	November	71.98%	Rainfall Mar. (100) | Max. temp. Feb. (93), Mar. (92)
	December	71.12%	Rainfall Jan. (100), Nov. (76), Apr. (71)
	January	72.41%	Rainfall Apr. (100), Mar. (93) | Feb. (83)
	February	56.90%*	Rainfall July (100) | Dist. urban (75) | Rainfall May (61)
	March	**78.45%**	**HeatU. Summer (100) | Max. temp. Mar. (96), Aug. (91)**
	April	**76.72%**	**Min. elev. (100) | Min. temp. May (72), July (70)**
	May	75.86%	Rainfall May (100), July (99), Dec. (89)
2012/2013	Seasonal	72.41%	HeatU. Summer (100), ann. (68), Winter (62)
	November	**78.88%**	**Heavy frost occ. (100) | Rainfall Jan. (85) | HeatU. ann. (76)**
	December	68.10%	Rainfall May (100) | Dist. grapes (93) | Rainfall July (84)
	January	64.66%	Min. temp. June (100), July (99), Aug. (72)
	February	**76.72%**	**Rainfall ann. (100), Sept. (75), May (66)**
	March	72.84%	Min. temp. June (100) | Rainfall ann. (96) | Min. temp. Aug. (87)
	April	71.41%	Rainfall Jan. (100), May. (82), June (69)
	May	68.01%	Rainfall May (100), June (94) | HeatU. Summer (91)
2013/2014	Seasonal	73.71%	Rainfall July (100), May (97), Jan. (78)
	November	61.21%	Rainfall Feb. (100| Min. temp. May (45) | Min. elev. (45)
	December	62.93%	Rainfall July (100), May (83), June (64)
	January	**75.00%**	**Min. elev. (100) | Rainfall Feb. (96) | Heavy frost occ. (82)**
	February	40.09%*	Rainfall Nov. (100) | HeatU. ann. (83) | Positive chill units (80)
	March	68.10%	Rainfall June (100), May (77), Mar. (70)
	April	**77.59%**	**Rainfall Jan. (100), Sept. (44), ann. (38)**
	May	73.71%	Rainfall June (100), May (93), Dec. (90)

Chapter 5

Investigating the robustness of the random forest machine learning algorithm to classify *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) population hot- and cold spots in two fruit producing regions of the Western Cape, South Africa.

Abstract

Understanding which factors drive the spatial distribution of tephritid fruit fly pests on a regional scale can help improve area-wide management programmes in terms of planning and application of management actions. The random forest (RF) classifier has the ability to model complex relationships in ecological datasets. However, such models need to be evaluated in terms of their robustness to handle complex problems. In this study, two area-wide *Ceratitis capitata* (Diptera: Tephritidae) trapping datasets, collected in two geographically separated regions, were used to quantify the spatial distribution of the trapping data into hot- and cold spots (HCSs) using spatial analysis. RF, using geographic variables, were used to model the observed HCSs. The spatial analysis results from the two regions were combined and the RF model classification accuracies of the combined regions were compared to the results from the individual regions, to investigate the robustness of the RF algorithm in modelling complex datasets. The results showed that the combined model accuracies were not significantly lower than those of the individual regions (but in some cases did significantly increase). The drivers of *C. capitata* spatial distribution were different between regions, but distance to urban areas emerged as a strong driver early in the fruiting season in all scenarios. The findings show that RF is a useful tool for investigating the spatial distribution of area-wide tephritid fruit fly trapping data, and that it can handle complex classification problems. The area-specific RF models provided invaluable information, which could be used to improve the planning and implementation of area-wide management programmes in heterogeneous agricultural landscapes.

Introduction

About 35% of fruit fly (Diptera: Tephritidae) species attack soft fruits, some of these species are major pests of commercial fruits globally (White & Elson-Harris 1992) and in South Africa (see Karsten *et al.* 2018). Understanding the relationships between the spatial distribution of these flies and their biotic and abiotic environment (Savopoulou-Soultani *et al.* 2012; DeVilliers *et al.* 2013; Flores *et al.* 2016), will lead to more focussed, effective management and cost savings (Nestel *et al.* 2002; Papadopoulos *et al.* 2003; Sciarretta & Trematerra 2011).

The current chapter expands on Chapter 4, in which long-term area-wide *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) trapping data from the Elgin, Grabouw, Villiersdorp and Vyeboom (EGVV) region, were subjected to spatial analysis to determine *C. capitata* long-term spatial patterns. The trap catches were characterised as hot- and cold spots (HCSs), where hot spots are indicative of clustering of high trap catches in space, while cold spots are indicative of clustering low trap catches in space. The main geographic drivers of these HCSs were identified, using the random forest (RF) classification algorithm, a machine learning (ML) algorithm (Breiman 2001), stemming from classification and regression trees (CART's) (Breiman *et al.* 1984; Breiman 2001). The results obtained from Chapter 4 were used in the analysis of this chapter, while data from a new study area, the Warmbokkeveld (WB), was also introduced.

The WB has been under an area-wide fruit fly management programme since 2010 (Barnes *et al.* 2015). It is a geographically isolated region, where the commercial production of different varieties and cultivars of pome and stone fruit is the major agricultural practice. Similar to EGVV, home gardens (farm and urban) and agricultural lands in WB are primary targets for fruit fly management.

In heterogeneous agricultural regions, such as EGVV and WB, host availability, climate, geography and landscape structure may vary, resulting in differences in their suitability for fruit fly populations (Papadopoulos *et al.* 2003; Puche *et al.* 2005; Sciarretta & Trematerra 2011). This can lead to differences in the drivers of fruit fly population HCSs in different regions. Thus the ecological relationships shaping these HCSs are complex (Mendelsohn *et al.* 2018). However, a number of authors have shown that the RF algorithm is robust and is able to model such complex relationships in ecological data (Guisan & Zimmermann 2000; Cutler *et al.* 2007; Olden *et al.* 2008; Elith & Leathwick 2009; Ruiz *et al.* 2010).

The aim of this chapter was to investigate the overall RF model accuracies and the most important drivers from WB and EGVV as well as a combined dataset (WB-EGVV), in relation to one another, to determine the robustness of the RF models when the sample size and variability in the predictor variables are increased. This was achieved, firstly, by determining the spatial patterns of *C. capitata* in WB, and identifying the most important geographic drivers of *C. capitata* in WB, using RF classification and secondly combining the spatial pattern data for WB and EGVV, and conducting RF classification analysis on the WB-EGVV dataset.

Warmbokkeveld background and data collection

The WB is situated around the towns of Ceres (-33.368835°, 19.311700°) and Prince Alfred Hamlet (-33.286263°, 19.326508°), which is a deciduous fruit producing region. Fruit production in the WB is mostly centered around the towns of Ceres and Prince Alfred Hamlet, which are in close proximity to one another (± 8,5 km), with fruit orchards in close proximity to urban areas. Farm and urban home gardens are dispersed amongst commercial fruit orchards. The total area covered by fruit orchards in WB is 4880 ha, consisting of pome

fruit (2944 ha), stone fruit (1888 ha) and wine grapes (48 ha) (Western Cape Department of Agriculture 2014) (Figure 1). The fruiting season in the WB starts in October/November and ends in May/June with most stone fruit cultivars ripening first while the last fruit to ripen are pome fruit cultivars. The general harvest period in WB is similar to the harvest period in EGVV (Manrakhan & Addison 2014), however, one to four week differences in harvest times might occur between the two regions, due to differences in climatic factors (i.e. heat units) as well as fruit cultivar and varietal differences.

In addition to fruit farming, other land uses in WB, include agricultural production facilities (fruit and potato packhouses), grain production, vegetable production, chicken hatcheries, dairies as well as natural vegetation. Alternate fruit fly host plants recorded in urban and farm home gardens in the region include *Citrus* spp., loquat (*Eriobotrya japonica* (Thunb.) Lindley), kei-apple (*Dovyalis caffra* (J.D. Hook & Harvey) J.D. Hook), guava (*Psidium guajava* L.) and fig (*Ficus carica* L.) (Manrakhan & Addison 2014).

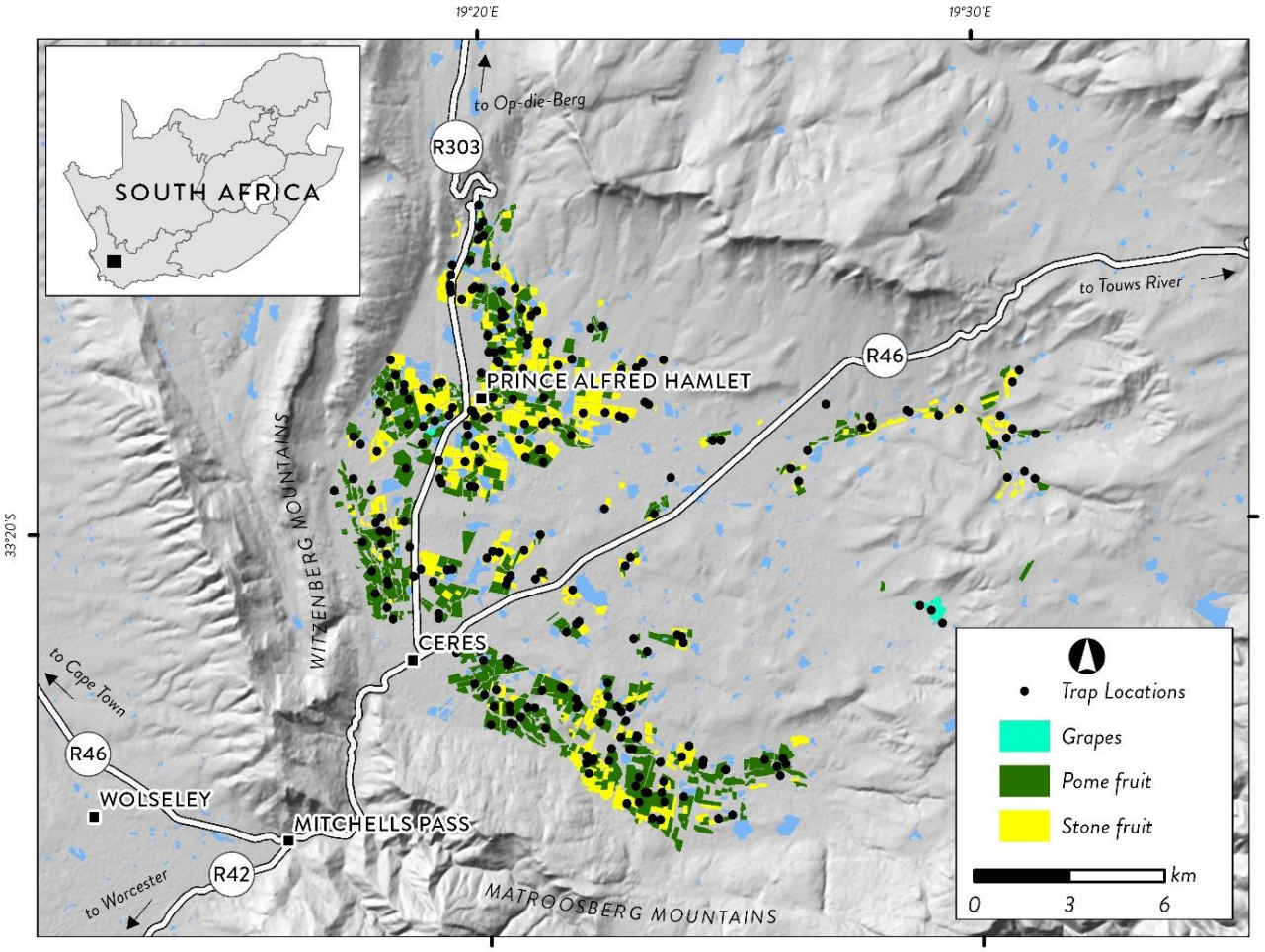

Figure 1. The distribution of *Ceratitis capitata* trap locations and the main fruit types planted in the Warmbokkeveld, based on the Western Cape Department of Agriculture's 2013 Flyover crop census (Western Cape Department of Agriculture 2014).

WB has a Mediterranean climate with hot, dry summers and cold, wet winters, similar to the climate observed in EGVV. However, some topographic and climatic differences do occur between the two regions, especially in the value ranges of different topographic and climatic variables (Table 1). WB has a smaller elevation range (467 m), compared to EGVV (976 m) but the regions are similar in terms of slope. There are differences in the long-term climatic variables between the two regions: The WB has more extreme temperatures with the long-term mean minimum temperature of the coldest month being 3°C colder than EGVV. WB is also drier with more chill units, less heat units and more heavy frost occurrences than EGVV. WB is situated at the foot of the Matroosberg and Witzenberg mountains, which receive

frequent snowfall in winter. The coldest month in WB and EGVV is July, while the hottest month in both regions is January (Schulze 2006).

Table 1. The minimum and maximum values and the range of the values of topographic and long-term (30 years) climatic variables of two distinct deciduous fruit growing regions, the Warmbokkeveld (WB) and the Elgin/Grabouw/Villiersdorp/Vyeboom (EGVV) region and their combined variable values.

Variable	Minimum and maximum value			Value range		
	WB	EGVV	Combined	WB	EGVV	Combined
Elevation (m)	447 - 914	104 - 1080	104 - 1080	467	976	976
Slope gradient (°)	0 - 45	0 - 46	0 - 46	45	46	46
Mean annual temperature (°C)	14 - 16	13 - 16	13 - 16	2	3	3
Mean annual maximum temperature (°C)	21 - 23	19 - 22	19 - 23	2	3	4
Mean annual minimum temperature (°C)	7 – 9	8 – 11	7 – 11	2	3	4
Mean maximum temperature of hottest month (°C)	27 - 29	24 - 28	24 - 29	2	4	5
Mean maximum temperature of coldest month (°C)	15 - 16	14 - 17	14 - 17	1	3	3
Mean minimum temperature of hottest month (°C)	12 - 14	12 - 15	12 - 15	2	3	3
Mean minimum temperature of coldest month (°C)	2 - 4	5 - 7	2 - 7	2	2	5
Mean annual rainfall (mm)	467 - 696	676 - 1032	467 - 1032	229	356	565
Mean rainfall of wettest month (mm)	75 - 118	104 - 172	75 - 172	43	68	97
Mean rainfall of driest month (mm)	9 - 14	18 - 32	9 - 32	5	14	23
Mean accumulated positive chill units	1184 - 1785	651 - 1438	651 - 1785	601	787	1134
Mean number of heavy frost occurrences	14 - 33	0 - 8	0 - 33	19	8	33
Mean annual heat units	1333 - 2103	1628 - 2410	1333 - 2410	770	782	1077
Mean summer heat units	1142 - 1683	1267 - 1706	1142 - 1706	541	439	564
Mean winter heat units	200 - 423	360 - 716	200 - 716	223	356	516

The area-wide fruit fly management programme under which WB falls, incorporates routine fruit fly monitoring, ground releases of sterile *C. capitata* male flies in urban areas and the coordination of aerial bait application technique (BAT) in orchards (Barnes *et al.* 2015), making it an area-wide integrated pest management (AW-IPM) programme. BAT involves the application of a protein and pesticide mixture, which is applied aerially before fruits are harvested to prevent crop losses and damage to fruit destined for export (Barnes *et al.* 2015). Four aerial BAT applications were made between January and March 2016 (pers.

comm. Nando Baard, FruitFly Africa). Sterile *C. capitata* males were released on the ground in urban areas on a weekly basis.

The dataset that was used for the analysis in this chapter was sourced from the FruitFlyAfrica (Pty) Ltd. centralised database for WB. *C. capitata* trap catch data for the 2015/2016 fruiting season (hereafter season), were sourced from this database. At the time of the analysis, this was the only available trapping data for WB, and extended from week 45 of 2015 to week 22 of 2016. Traps were distributed among commercial fruit orchards and home gardens (Figure 1), according to the trapping guidelines of the International Atomic Energy Agency (IAEA) (IAEA 2003). Trapping data were obtained from 246 Chempac© Yellow Bucket traps (McPhail-type trap) with Vapona strips (dichlorvos) as retention system baited with three-component Biolure® Fruit Fly (Chempac Pty. Ltd). Biolure® is a synthetic food-based lure (Ammonium acetate 211g/kg; Trimethylamine hydrochloride 91 g/kg; 1,4-diaminobutane (Putrescine) 3g/kg). It attracts both male and female *Ceratitis* spp., but is biased toward females (Heath *et al.* 1997; Epsky *et al.* 1999; Miranda *et al.* 2001; Leza *et al.* 2008). This trapping system is recommended in *C. capitata* AW-IPM programmes where SIT is incorporated, in order to attract more wild females and fewer sterile and wild males (IAEA 2003; Ekesi *et al.* 2005). Data from 16 Chempac© Yellow Delta traps with sticky pads as retention system, baited with Chempac Fruit Fly Lure (Chempac Pty. Ltd) with the male attractant trimedlure as active ingredient (1,0 g), *tert*-butyl 4 (and 5)-chloro- trans-2-methylcyclohexane-1-carboxylate, were also included in the analysis. No sterile fruit fly trap data were recorded in the 2015/2016 in WB, as no sterile males were released in commercial orchards on a regular basis. Thus, no sterile-to-wild ratio (SWR) could be incorporated into the analysis.

Data analysis

For data processing, spatial analysis and ML procedures methods of WB data, refer to the data processing, spatial analysis and ML sections in Chapter 4. Viable traps were selected, after which ordinary kriging (OK) (Childs 2004) was performed to incorporate the relative location of each trap into the analysis (Midgarden *et al.* 2014). OK was used to spatially interpolate and generalize trap catch data into 100 m resolution continuous raster layers. Continuous raster layers were created for monthly and seasonal trap catch data. The 100 x 100 m raster layers were then further aggregated into 1 km^2 zones using zonal statistics. Similar procedures were conducted for a range of predictor variables (Table 2), considered in the ML analysis. The Euclidian distance from different crop types and urban areas to each zone was also calculated to investigate the effect of these landscape factors on the spatial distribution of *C. capitata* trap catch.

The total number of traps in the WB trapping dataset used in the analysis were 262. The final 1 km^2 zone grid used for the spatial analysis contained at least one fruit fly trap in each zone, resulting in a total of 118 zones considered for the analysis.

The optimised hot spot analysis in ArcGIS 10.5 was used, to quantify the spatial distribution *C. capitata* trap catch. This tool classifies statistically significant spatial clusters of high values (high trap catch/hot spots) and low values (low trap catch/cold spots) relative to the entire study region. The tool uses the Getis-Ord Gi* statistic to identify statistically significant clusters of high and low values (Getis & Ord 1992b; Ord & Getis 1995). Zones were binned into three classes/groups, reflecting HCSs at 99, 95 and 90 % confidence levels, while non-significant zones were also binned (see the Spatial analysis section in Chapter 4).

RF was used to determine the main drivers of HCSs in WB, using a set of predictor variables (see Table 2). For an explanation of RF, refer back to the ML section in Chapter 4. The Gini method was used to generate the variable importance lists (VIL's) for each RF model. This method generates a variable importance score, which is an indication of how often a predictor variable were used to split the data at each node and it is also a measure of the discriminative value of each predictor variable in classifying the data (Menze *et al.* 2009).

The same data analysis procedures employed for WB and EGVV (see chapter 4) were carried out on a combined dataset of WB and EGVV (WB-EGVV). Minor differences occurred in the number of landscape predictor variables used in the combined data analysis, due to differences in crop and landscape composition between the two regions (Table 2). The following landscape variables were excluded from the combined data analysis: distance to home gardens, distance to pome fruit, distance to citrus and distance to berries.

All spatial analyses and map production were done in ArcGIS 10.5 software, while the RF modeling was done using SALFORD PREDICTIVE MODELER® software (www.salford-systems.com).

To compare the frequency of misclassifications and correct classifications, seasonal and monthly RF model classification accuracies were compared between the three different datasets (WB, EGVV and WB-EGVV combined), using a Chi-square test of independence. To provide for data that are not independent from one another, the Rao and Scott adjustment on the Chi-square test was selected (Rao & Scott 1987). The Chi-square analyses were carried out using STATISTICA version 13.3 (TIBCO Software Inc 2017), while the Rao & Scott adjustment were run using the 'survey' package for R v.3.5.1 (R Core Team 2018).

Table 2. Predictor variables used in the random forest models in the Warmbokkeveld (WB) Elgin/Grabouw/Villiersdorp/Vyeboom (EGVV) and the two regions combined (WB-EGVV). The differences in predictor variables between the two regions are in bold.

Variables	WB	EGVV	WB-EGVV	Num. of variables			Temporal scale
				WB	EGVV	WB-EGVV	
Climate	Min. temp.	Min. temp.	Min. temp.	26	26	26	Ann. and monthly
	Max. temp.	Max. temp.	Max. temp.	26	26	26	Ann. and monthly
	Mean temp.	Mean temp.	Mean temp.	26	26	26	Ann.
	Rainfall	Rainfall	Rainfall	26	26	26	Ann. and monthly
	Acc. positive chill units	Acc. positive chill units	Acc. positive chill units	1	1	1	May to Sept.
	Mean number of heavy frost occurrences	Mean number of heavy frost occurrences	Mean number of heavy frost occurrences	1	1	1	Ann.
	Acc. heat units (base 10°C)	Acc. heat units (base 10°C)	Acc. heat units (base 10°C)	3	3	3	Ann., summer (Oct. to Mar.) and winter (Apr. to Sept.)
Topographic	Elevation	Elevation	Elevation	5	5	5	n/a
	Slope	Slope	Slope	5	5	5	n/a
Landscape	Dist. urban	Dist. urban	Dist. urban	1	1	1	n/a
	Dist. home gardens	**Dist. citrus**	**N/A**	1	1	0	n/a
	Dist. stone fruit	Dist. stone fruit	Dist. stone fruit	1	1	1	n/a
	Dist. pome fruit	**Dist. berries**	**N/A**	1	1	0	n/a
	Dist. grapes	Dist. grapes	Dist. grapes	1	1	0	n/a

Results

The mean number of *C. capitata* catches per trap per week (FTW) for both regions show population peaks during March-May of each season, but also when all seasons were combined in EGVV (Figure 2). High numbers of FTW were observed in WB, as well as EGVV (2010/2011 season), in particular.

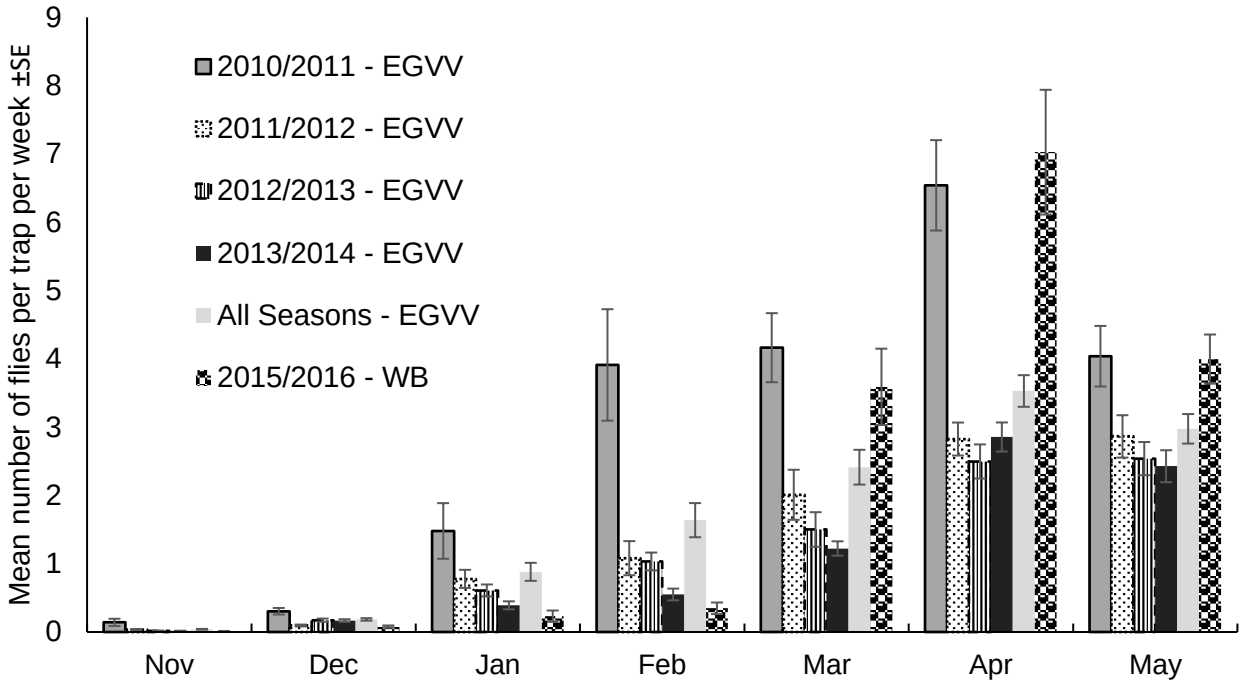

Figure 2. The monthly mean (±SE) number of *Ceratitis capitata* adults per trap per week in Elgin/Grabouw/Villiersdorp/Vyeboom (EGVV) (2010/2011-2013/2014) and the Warmbokkeveld (WB) (2015/2016).

The spatial analysis revealed seasonal (2015/2016) *C. capitata* hot spots just north of Ceres (in close proximity to each other), with scattered cold spots south of Ceres and to the north and west of Prince Alfred Hamlet (Figure 3h). Monthly *C. capitata* population hot spots (Figure 3a-g) generally occurred between Ceres and Prins Alfred Hamlet in close proximity to the towns. There were some exceptions where hot spots also occurred east of the region in December (Figure 3b) and March (Figure 3e) and south of Ceres in January (Figure 3c) and May (Figure 3g). Generally, however, hot spots occurred in close proximity to the urban areas. Cold spots were mainly scattered across WB, except for March (Figure 3e) and April (Figure 3f) when cold spots were concentrated to the south of Ceres.

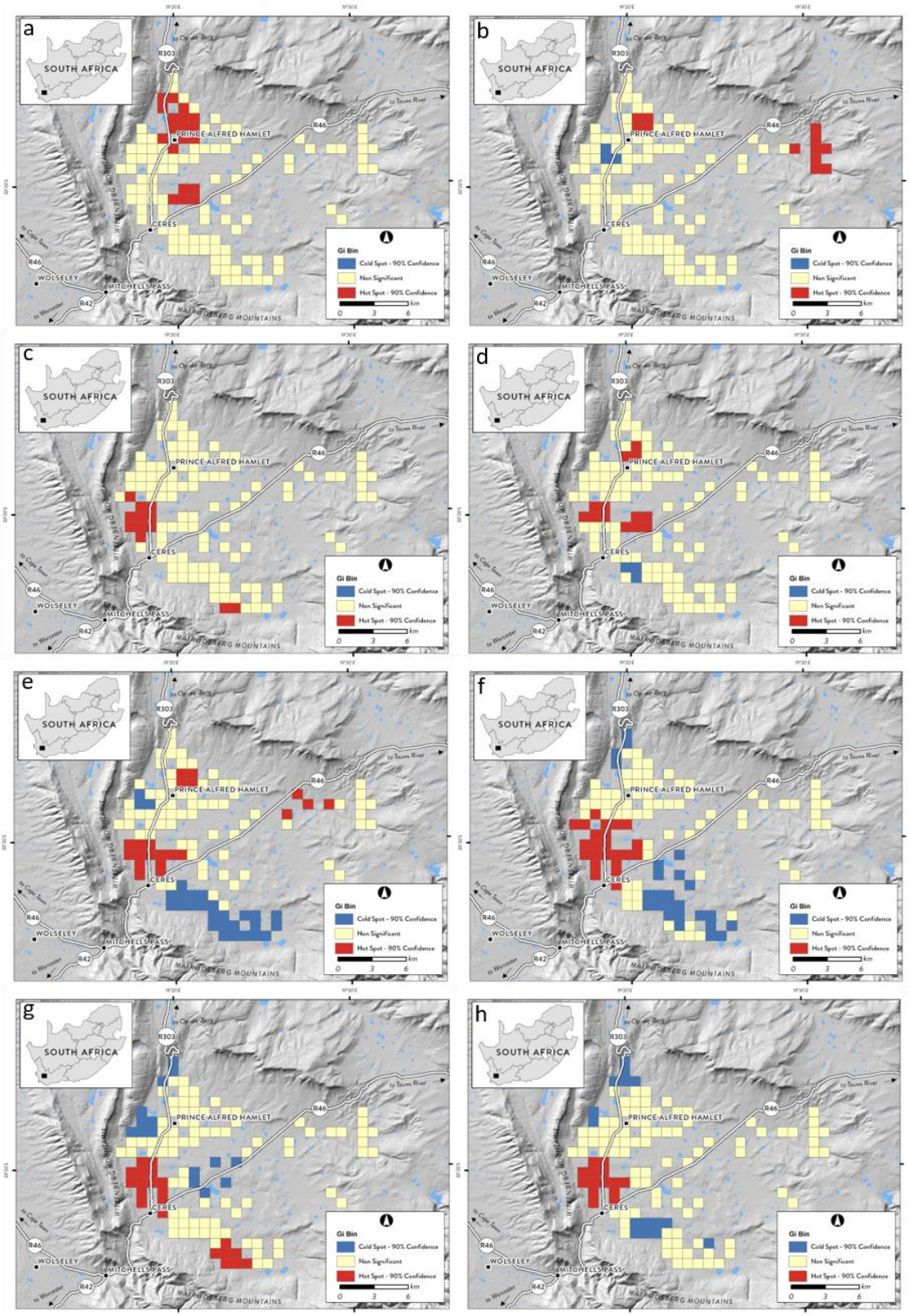

Figure 3. *Ceratitis capitata* hot- and cold spots for the 2015/2016 season in the Warmbokkeveld for (a) November; (b) December; (c) January; (d) February; (e) March; (f) April; (g) May and for the (h) entire season classified. Red zones indicate hot spots; blue zones indicate cold spots while yellow zones indicate no significant spatial pattern.

The RF models performed well in classifying *C. capitata* HCSs in WB. The classification accuracies of the March and April HCSs, as well as the seasonal HCSs were all above 70%, with the April model yielding the best result (79.66 %) (Table 3). In WB, temperature related predictor variables, emerged as the strongest predictors in the majority of the models, including summer heat units in the seasonal model, while the most important predictor variable in the March model was long-term rainfall. Distance to urban areas emerged as the most important variable in November, however, the overall model accuracy was lower (64.41%) compared to the other models, especially the March (74.58%) and April (79.66%) models.

Table 3. Summary of the random forest classification results of *Ceratitis capitata* spatial patterns Warmbokkeveld (2015/2016), EGVV (all seasons combined) and the Warmbokkeveld and EGVV combined. Variable importance lists were calculated using the Gini method. Bold text indicates the two most accurate models in the region while (*) indicates class imbalances in the target variable.

Region	Time period	Overall model accuracy (%)	Most important variables in model (importance score)
Warmbokkeveld (WB)	**Seasonal**	**77.97%**	**HeatU. summer (100) | HeatU. ann. (77) | Mean elev. (74)**
	November	64.41% *	Dist. urban (100) | Rainfall Sept. (89), Aug. (79)
	December	64.41%	Dist. grapes (100) | Stdev. Rainfall Mar. (88) | Stdev. max. temp. May (75)
	January	69.49% *	Min. temp. ann. (100), Nov. (89), Apr. (84)
	February	66.10%	Min. temp. Aug. (100) | Min. elev. (86) | HeatU. winter (76)
	March	74.58%	Rainfall Mar. (100) | Max. temp. July (87) | Min. temp. July (72)
	April	**79.66%**	**HeatU. summer (100) | Min. temp. Dec. (91) | Min. elev. (86)**
	May	65.25%	Min. temp. ann. (100), Apr. (98), May (90)
EGVV	Seasonal	66.38%	Rainfall Dec. (100), ann. (77), May (73)
	November	53.02%*	Dist. urban (100) | Rainfall Mar. (73), Dec. (67)
	December	65.95%	Rainfall Jan. (100), Nov. (89), May (87)
	January	68.10%	Rainfall Mar. (100), Feb. (88) | Min. temp. May (60)
	February	52.59%*	Rainfall Oct. (100) | Max. temp. Aug. (81) | Rainfall Apr. (75)
	March	**79.74%**	**Rainfall Dec. (100), June (89), ann. (89)**
	April	**79.31%**	**Max. temp. Nov. (100), ann. (89) | Rainfall Mar. (87)**
	May	69.83%	Rainfall Nov. (100), ann. (99), May (86)
Warmbokkeveld and EGVV combined	Seasonal	70.86%	Rainfall ann. (100), Dec. (99) | Min. temp. Feb. (85)
	November	51.71%*	Dist. urban (100) | Rainfall July (40), June (35)
	December	68.86%	Rainfall Jan. (100) | Max. temp. Sept. (89) | Rainfall Dec. (85)
	January	72.00%	Rainfall Mar. (100), Feb. (87) | Min. temp. June (81)
	February	63.43%	Max. temp. Jan. (100), May (96), Sept. (95)
	March	**78.86%**	**Rainfall ann. (100), Dec. (79), Sept. (79)**
	April	**81.14%**	**Rainfall Mar. (100), Dec. (72), ann. (55)**
	May	67.14%	Mean temp. ann. (100) | Rainfall ann. (88), Sept. (80)

The only significant differences in model accuracies between the three datasets were observed in the November ($p = 0.05$) and February ($p = 0.01$) models. In November, the RF model in WB (64.41%) outperformed the models in EGVV (53.02%) and WB-EGVV (51.71%). Similarly, the WB February model (66.10%) outperformed the WB-EGVV

(63.43%) and the EGVV (52.59%) models. There were no significant differences (p = > 0.05) in the accuracies of the other models between the three datasets.

Long-term rainfall generally emerged as the strongest predictor of *C. capitata* HCSs in EGVV, while in WB, temperature related predictors were the most common drivers of trap catch HCSs (Table 3). However, rainfall emerged as the most important predictor in the March model of both regions. Distance to urban areas emerged as the most important variable of the November RF model, in all scenarios. Model accuracies between the three datasets were similar, in that the models performed the best in classifying HCSs in March and April, except for in WB, where the seasonal model was the second most accurate model, after the April model. In WB-EGVV, rainfall emerged as the dominant predictor variable, with temperature only emerging as the most important variable in the February and May models. This is most similar to the VIL results obtained from the EGVV models.

Discussion

Unmanaged winter hosts in urban home gardens can act as a breeding ground for *C. capitata* populations when no hosts are available in surrounding commercial orchards, from where they can populate commercial orchards when fruits become available and susceptible for infestation (Myburgh 1956; Israely *et al.* 1997; Barnes 2008; DeVilliers *et al.* 2013; Manrakhan & Addison 2014), also see Chapter 2. A visual interpretation of the OHA maps suggests that this movement is taking place in WB, with hot spots occurring in close proximity to urban areas. The RF results support this, in that distance to urban areas emerged as the most important predictor variable in the November models, of all three datasets (WB, EGVV and WB-EGVV), although the model accuracies were relatively low. This data is, therefore, able to quantify the importance of these alternate hosts through the variable, ‘distance to urban areas’, which supports suggestions made by Barnes and Venter

(2006), who indicated that flies from urban home garden winter hosts move into nearby commercial orchards, when fruits become susceptible for oviposition. The RF modelling also suggests that hosts in urban areas only drive/contribute to the spatial distribution of *C. capitata* populations during the early fruiting season (November). An explanation for this could be that the flies breed and disperse from commercial orchards once the first population that has originated in urban home garden winter hosts, has completed a full life cycle. Thus, the proximity to urban areas loses its importance as a driver of *C. capitata* population hot spots in commercial orchards during the rest of the season and other factors, such as climate and ripening commercial hosts, become stronger drivers of HCSs. It is suggested that management resources should be more focussed on urban areas and farm gardens in the period leading up to November. Thereafter, resources should be directed more toward commercial orchards in order to suppress the breeding populations there. Such a precision approach to managing fruit flies on a regional basis will contribute to a more effective management strategy and improve efficiency and sustainability of AW-IPM programmes.

The differences in the RF VIL's between WB and EGVV are indicative that the driving factors of *C. capitata* HCSs are area-specific. *C. capitata* populations have been shown to respond spatially to changes in the environment (climate) and the variability that exists in the type of host species, the ripening sequence of fruit as well as fruit availability within regions (Nestel *et al.* 2002; Papadopoulos *et al.* 2003; Sciarretta & Trematerra 2011). Therefore, the range differences in the minimum and maximum values of the topographic and long-term (30 years) climatic variables and possible host differences between the two regions, all contributed to the difference in the drivers of *C. capitata* HCSs. One might also argue that the reason why temperature related variables were the most dominant drivers of *C. capitata* HCSs in WB and not EGVV, where rainfall was more important was because of the climatic differences between the regions. WB is generally drier and hotter than EGVV, but the

minimum and maximum values of the long-term temperature related variables in WB are more variable compared to EGVV. *Ceratitis capitata* prefers hot and dry climates compared to cold and wet climates (Duyck & Quilici 2002; Duyck *et al.* 2006b; Nyamukondiwa & Terblanche 2009). Therefore, the differences in the variability of the long-term climatic variables, between the regions, may have influenced the differences in the main drivers of *C. capitata* population HCSs, in the different regions. It is clear from the RF modelling results, that it is a complex combination of different factors driving *C. capitata* HCSs on a regional basis. Furthermore, the way in which the flies react to their environment is dynamic, which is portrayed by the differences seen in the spatial distribution of HCSs, between months and even between seasons.

The RF model accuracies from WB-EGVV indicated that increasing the complexity of the classification problem, did not have a significant negative influence on model accuracies. In contrast, the model accuracies significantly improved, in some cases. This is encouraging as it suggests that the RF ML algorithm could handle these complex ecological interactions.

In all the RF models for WB, EGVV and WB-EGVV, the March and April models emerged as the strongest. This is attributed to the fact that in these months (autumn) fruit fly numbers are at a peak, after which population numbers started to decrease (see also De Villiers *et al.* 2013; Manrakhan and Addison 2014). Mavrikakis *et al.* (2000) suggested that the decrease in trap catches during winter and early spring in Southern Greece, which also has a Mediterranean climate, might be due to lower temperatures causing some adult mortality, decrease in activity as well as low trap efficiency. However, the drop in fruit fly numbers could also be influenced by host availability (Sciarretta & Trematerra 2011). During March and April, low and high trap catches were more clustered, indicating that the flies are more aggregated during this period, perhaps because of some commercial orchards still having

fruit available for the flies to use as oviposition sites, around which they clustered. This suggests that the RF classification algorithm, in the context of this study, yielded the highest model accuracies when *C. capitata* populations are most clustered in space.

Conclusion

The robustness of the RF classifier to classify *C. capitata* HCSs on a regional scale was investigated by increasing the sample size and the variability in the predictor variables. The results from this study provided new evidence of the importance of long-term environmental variables in determining the spatial distribution of fruit flies. Drivers of *C. capitata* HCSs were different between the regions, because of climatic and geographic differences between the regions. Therefore, it is recommended to use regional-specific models. This study provides and adds valuable information regarding the drivers of *C. capitata* HCSs in heterogeneous fruit production systems. This information will assist the deciduous fruit industry in improving its management of this global economically important pest, by using the information in the planning and strategic decision making stages of area-wide fruit fly management programmes. It also paves the way for future research on other important fruit flies (*Bactrocera dorsalis*) and moth pests (*Thaumatotibia leucotreta)* which are currently threatening the deciduous fruit industry in South Africa.

This study further shows that RF classification is robust and powerful tool to investigate and explain complex relationships often found in environmental and ecological datasets. To investigate the full extent of the driving force of hosts on *C. capitata* HCSs, area-wide cultivar information should be incorporated to make the models more specific, which may also improve our understanding of the effect of hosts on HCSs at a regional scale. This chapter demonstrates that the results of Chapter 4 are not universal (i.e. rainfall did not always emerge as the most important predictor variable) and the methodology developed and

demonstrated in Chapter 4, can easily be extended to other (and larger) areas. This is indicative that these RF classification models are universally applicable and that scale of analysis is important when area-wide fruit fly managers interpret *C. capitata* HCSs.

References

BARNES, B. 2008. Why have fruit flies become so troublesome? *South African Fruit Journal* **7**(4): 56–57.

BARNES, B.N., HOFMEYR, J.H., GROENEWALD, S., CONLONG, D.E. & WOHLFARTER, M. 2015. The sterile insect technique in agricultural crops in South Africa : a metamorphosis .... but will it fly ? *African Entomology* **23**(1): 1–18.

BARNES, B.N. & VENTER, J.H. 2006. The South African Fruit Fly Action Plan – Area-wide Suppression and Exotic Species Surveillance. *Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance*. 271–283. Salvador, Brazil.

BREIMAN, L. 2001. Random forests. *Machine Learning* **45**(1): 5–32.

BREIMAN, L., FRIEDMAN, J., OLSHEN, R. & STONE, C. 1984. *Classification and Regression Trees*. Wadsworth, Belmont, California.

CHILDS, C. 2004. Interpolating surfaces in ArcGIS spatial analyst. *ArcUser, July-September* **3235**: 569.

CUTLER, D.R., EDWARDS, T.C., BEARD, K.H., CULTLER, A., HESS, K.T., GIBSON, J. & LAWLER, J.J. 2007. Random Forests for Classification in Ecology. *Ecological Society of America* **88**(11): 2783–2792.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

DUYCK, P.F. & QUILICI, S. 2002. Survival and development of different life stages of three *Ceratitis* spp. (Diptera : Tephritidae) reared at five constant temperatures. *Bulletin of Entomological Research* **92**(6): 461–469.

EKESI, S., LUX, S.A. & BILLAH, M. 2005. Field comparison of food-based synthetic attractants and traps for African Tephritid fruit flies. *Development of Improved Attractants and Their Integration into Fruit Fly SIT Management Programmes*. 205–230. International Atomic Energy Agency, Vienna.

ELITH, J. & LEATHWICK, J.R. 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics* **40**(1): 677–697.

EPSKY, N.D., HENDRICHS, J., KATSOYANNOS, B.I., VASQUEZ, L.A., ROS, J.P., ZUMREOGLU, A., PEREIRA, R., BAKRI, A., SEEWOORUTHUN, S.I. & HEATH, R.R. 1999. Field Evaluation of Female-Targeted Trapping Systems for *Ceratitis capitata* (Diptera:Tephritidae) in Seven Countries. *Journal of Economic Entomology* **92**: 156–164.

FLORES, S., MONTOYA, P., RUIZ-MONTOYA, L., VILLASEÑOR, A., VALLE, A., ENKERLIN, W. & LIEDO, P. 2016. Population Fluctuation of *Ceratitis capitata* (Diptera: Tephritidae) as a Function of Altitude in Eastern Guatemala. *Environmental Entomology* **45**(4): 802–811.

GETIS, A. & ORD, J.K. 1992. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis* **24**(3): 189–206.

GUISAN, A. & ZIMMERMANN, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.

HEATH, R.R., EPSKY, N.D., DUEBEN, B.D., RIZZO, J. & JERONIMO, F. 1997. Adding methyl-substituted ammonia derivatives to a food-based synthetic attractant on capture of the Mediterranean and Mexican fruit flies (Diptera: tephritidae). *Journal of Economic Entomology* **90**(6): 1584–1589.

IAEA. 2003. Trapping guidelines for area-wide fruit fly programmes. International Atomic Energy Agency, Vienna.

ISRAELY, N., YUVAL, B., KITRON, U. & NESTEL, D. 1997. Population fluctuations of adult Mediterranean fruit flies (Diptera: Tephritidae) in a Mediterranean heterogeneous agricultural region. *Environmental Entomology* **26**(6): 1263–1269.

KARSTEN, M., BARTON, M., ADDISON, P., ADDISON, M.F. & TERBLANCHE, J.S. 2018. A synthesis for managing invasions and pest risks simultaneously for tephritid fruit flies in South Africa. *Entomologia Experimentalis et Applicata*(2): 1–13.

LEZA, M.M., JUAN, A., CAPLLONCH, M. & ALEMANY, A. 2008. Female-biased mass trapping vs . bait application techniques against the Mediterranean fruit fly, *Ceratitis capitata* (Dipt., Tephritidae). *Journal of Applied Entomology* **132**(9–10): 753–761.

MANRAKHAN, A. & ADDISON, P. 2014. Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science* **70**(4): 651–660.

MAVRIKAKIS, P.G., ECONOMOPOULOS, A.P. & CAREY, J.R. 2000. Continuous Winter Reproduction and Growth of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Heraklion, Crete, Southern Greece. *Environmental Entomology* **29**(6): 1180–1187.

MENDELSOHN, O., DAYAN, T., AIDLIN-HARARI, S., SILBERSTEIN, M., ORLOV, V. & BLANK, L. 2018. Mediterranean fruit fly subplot hot spots prediction by experts' experience. *Journal of Applied Entomology* **00**: 1–9.

MENZE, B.H., KELM, B.M., MASUCH, R., HIMMELREICH, U., BACHERT, P., PETRICH, W. & HAMPRECHT, F.A. 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* **10**(213): 1–16.

MIDGARDEN, D.G., LIRA, E. & SILVER, M. 2014. Spatial Analysis of Tephritid Fruit Fly Traps. In: Shelly, T., Epsky, N., Jang, E.B., Reyes-Floris, J. & Vargas, R. (Eds.) *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. 277–320. Springer, Dordrecht.

MIRANDA, M.A., ALONSO, R. & ALEMANY, A. 2001. Field evaluation of Medfly (Dipt., Tephritidae) female attractants in a Mediterranean agrosystem (Balearic Islands, Spain). *Journal of Applied Entomology* **125**(6): 333–339.

MYBURGH, A.C. 1956. Bionomics and control of the fruit flies, *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.), in the Western Cape. University of Stellenbosch.

NESTEL, D., KATSOYANNOS, B.I., NEMNY-LAVY, E., MENDEL, Z. & PAPADOPOULOS, N.T. 2002. Spatial analysis of Medfly populations in heterogeneous landscapes. *Proceedings of 6th International Fruit Fly Symposium*. 35–43. Stellenbosch, South Africa.

NYAMUKONDIWA, C. & TERBLANCHE, J.S. 2009. Thermal tolerance in adult Mediterranean and Natal fruit flies (*Ceratitis capitata* and *Ceratitis rosa*): Effects of age, gender and feeding status. *Journal of Thermal Biology* **34**(8): 406–414. Elsevier.

OLDEN, J.D., LAWLER, J.J. & POFF, N.L. 2008. Machine Learning Methods Without Tears: A Primer for Ecologists. *The Quarterly Review of Biology* **83**(2): 171–193.

ORD, J.K. & GETIS, A. 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* **27**(4): 286–306.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PUCHE, H., MIDGARDEN, D.G., OVALLE, O., KENDRA, P.E., EPSKY, N.D., RENDON, P. & HEATH, R.R. 2005. Effect of Elevation and Host Availability on Distribution of Sterile and Wild Mediterranean Fruit Flies (Diptera: Tephritidae). *Florida Entomologist* **88**: 83–90.

R CORE TEAM. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

RAO, J.N.K. & SCOTT, A.J. 1987. On Simple Adjustments to Chi-Square Tests with Sample Survey Data. *The Annals of Statistics* **15**(1): 385–397.

RUIZ, M.O., CHAVES, L.F., HAMER, G.L., SUN, T., BROWN, W.M., WALKER, E.D., HARAMIS, L., GOLDBERG, T.L. & KITRON, U.D. 2010. Local impact of temperature and precipitation on West Nile virus infection in Culex species mosquitoes in northeast Illinois, USA. *Parasites & Vectors* **3**(1): 19.

SAVOPOULOU-SOULTANI, M., PAPADOPOULOS, N.T., MILONAS, P. & MOYAL, P. 2012. Abiotic factors and insect abundance. *Psyche (New York)* **2012**: 2012–2014.

SCHULZE, R. 2006. *South African Atlas of Climatology and Agrohydrology*. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

TIBCO SOFTWARE INC. 2017. Statistica (data analysis software system).

WESTERN CAPE DEPARTMENT OF AGRICULTURE. 2014. The Mapping of Agricultural Commodity Production and Agricultural Infrastructure in the Western Cape.

WHITE, I.M. & ELSON-HARRIS, M.M. 1992. *Fruit flies of economic significance: their identification and bionomics*. CAB International/Australian Centre for International Agricultural Research, Wallingford, UK/Canberra.

Chapter 6

General discussion

Since *C. capitata* and *Ceratitis quilicii* females both cause fruit damage, but also hold a phytosanitary risk for fruit exports (DeVilliers *et al.* 2013), it was important to investigate their spatio-temporal distribution and interaction between one another in a heterogeneous orchard environment to improve the location and timing of management actions at orchard level. This was done on a smaller farm-scale, so as to 1) include *C. quilicii*, which is not currently being monitored by FFA's monitoring programme; and 2) establish potential spatio-temporal interactions between the two species, which could be valuable to inform management programmes. *C. capitata* females significantly aggregated more towards the end of the season, while *C. quilicii* females significantly aggregated more towards the beginning of the season, driven by their physiological response to climate. They were also spatially associated, indicating a niche overlap associated with ripening hosts, which was evident from fruit rearing data. It was interesting to note that no *C. quilicii* were reared from apple, which was dominated completely by *C. capitata*. This may be related to characteristics of the host fruit and/or phenology and should be confirmed with further host preference studies. The most important habitats where both species aggregated were home garden hosts, natural vegetation, and on the edge of the orchard facing these habitats. Unmanaged home garden hosts sustained both species throughout the season, while *C. quilicii* also found refuge in natural vegetation, comprising mostly tall, dense trees and shrubs bearing minimal fleshy fruits. The spatial patterns and associations observed indicate that dispersal from home garden and natural vegetation hosts is prevalent, confirming the importance of managing these habitats to prevent infestations from outside the orchard. This was the first study investigating the spatial patterns and associations between *C. capitata* and *C. quilicii.* Further experiments should be conducted across multiple seasons, to

determine whether the spatial patterns and associations are consistent, while micro-climates within the orchard can also be investigated as possible drivers of in-orchard spatial distribution of these flies. Experiments determining the natural hosts of *C. quilicii* in the Western Cape are also needed.

Limited knowledge is available on *C. capitata* spatio-temporal distributions in large heterogeneous fruit producing regions. After conducting a small scale study on the spatial distribution of *C. capitata* (Chapter 2), it was necessary to increase the scale of analysis to provide similar information on an area-wide scale. Long-term (four seasons) area-wide *C. capitata* trapping data were sourced from FFA's AW-IPM programme in three sub-regions, Elgin/Grabouw, Villiersdorp and Vyeboom, collectively known as the EGVV region, which is a heterogeneous fruit-producing region. According to Cox and Vreysen (2005), any successful AW-IPM programme requires a clear understanding of the spatio-temporal distribution of the target pest in relation to its environment. Consequently, the spatio-temporal distribution of *C. capitata* populations were investigated by visualising fruit fly trap catches and symbolising the trap catches using subjective management thresholds, of more than seven flies per trap per week. As was to be expected, a certain amount of variation occurred in the number of flies caught between individual traps. Different management practices by individual growers would have had an effect on the varying trap catches, which could not be measured here due to a lack of AW-specific data. This variation can possibly also be ascribed to trap placement (Robacker *et al.* 1990), micro-climate, such as temperature, relative humidity and light intensity (see Drake 1994; Haniotakis 1974; Kapatos and Fletcher 1983; Kitron 1998) as well as the behaviour and physiological state of the flies at the time (Neuenschwander & Michelakis 1979; Vreysen & Saleh 2001; Coracini *et al.* 2004). The symbolising of the trap catches, indicated that a spatial trend was evident in EGVV, through visual inspection of the maps created, as traps with high fly catches occurred

more to the south-east of each region while traps with low trap catches occurred more to the north-west of each region, suggesting geographic factors to be influential. The high variability in trap catches between individual traps limited the quantification of spatial distribution patterns from which management-supporting recommendations could be made. Given that management decisions in AW-IPM programmes are generally taken at coarser spatial scales, it is recommended that the data is spatially and temporally aggregated into manageable units of approximately 1 km^2. This will allow for improved statistical quantifications and likely accentuate the spatial relationships between the fruit fly distributions and the geographic characteristics within the area.

The limitations in Chapter 3, as discussed above were addressed in Chapter 4. Identifying the underlying spatial processes which drive *C. capitata* spatial distribution in heterogeneous agricultural landscapes can improve management decisions in AW-IPM programmes (Midgarden *et al.* 2014). Using spatial analysis, the spatial distribution of *C. capitata* trap catches in EGVV (see Chapter 3) were quantified into zones where high (hot spots) and low (cold spots) trap catches cluster in space. Each zone was then statistically compared to a range of possible geographic drivers, making use of the random forest (RF) ML classifier, to identify the most important drivers of *C. capitata* hot- and cold spots (HCSs). Hot spots were concentrated in the hotter and drier south-eastern parts of each sub-region of EGVV, while cold spots were concentrated more towards the colder and wetter north-western parts, indicating a clear split in the spatial distribution of HCSs. This confirms the trend of high and low trap catches observed from Chapter 3. It is known that *C. capitata* is better adapted to drier and hotter climates (Duyck *et al.* 2006; Nyamukondiwa & Terblanche 2010). In La Réunion, *C. capitata* and *C. quilicii,* of which the latter prefers colder and wetter climates, segregate ecologically and geographically, based on niche-dependent competition and their physiological response to climate (Duyck *et al.* 2006). However, this could not be explored

on a larger scale due to the lack of *C. quilicii* trapping data, but the temporal segregation due to climate was evident from data obtained in Chapter 2 (early season vs late season spatial aggregation). Therefore, emphasis should be placed on monitoring for other economically-important fruit flies on an area-wide scale. Identifying the driving factors of the *C. capitata* HCSs using RF, indicated that management variables, represented by the sterile to wild male ratio (SWR), was not a significant driver of the long-term HCSs in EGVV. However, the functioning of SIT is largely based on the competition between sterile and wild male flies (McInnis *et al.* 1994); therefore, the effect of the SWR may have been concealed by the relatively coarse scale at which this study was conducted. Long-term HCSs in March and April, which are the peak months for *C. capitata* populations, were shown to be strongly driven by the long-term climate, especially rainfall of the preceding months. Therefore, the climate of the preceding months of the areas where March and April hot spots occurred, favoured *C. capitata* development, while the opposite was true for areas where cold spots occurred. Using long-term environmental data, the drivers of the spatial distribution of HCSs in EGVV could be explained, consequently improving our understanding of the role that stable geographic factors play in determining the spatial distribution of *C. capitata* populations in heterogeneous agricultural systems. This study presents area-wide fruit fly management programmes, especially those that incorporate SIT, with a tool to conduct more precise spatial planning, which could lead to better programme performance and reduced costs. A more proactive approach in assessing risk in terms of stable geographic characteristics of an area, rather than focusing on highly variable seasonal factors, are also presented. Future research should test whether the same geographic drivers of *C. capitata* HCSs emerge as being important in other fruit-producing regions of South Africa. The geospatial approach used in this study should provide a good foundation for such work. Furthermore, to the best of my knowledge, this was the first study investigating the effect of geographic environmental factors on the spatial distribution of tephritid fruit flies in a

heterogeneous fruit producing region, using ML and specifically RF. Area-wide data on host location, cultivar information, host phenology and data on management actions could be incorporated into the ML algorithms to make them more resilient. It is known that host type and host phenology play an important role in determining the spatial distribution of *C. capitata* (Katsoyannos *et al.* 1998; Papadopoulos *et al.* 2001, 2003; Vera *et al.* 2002; Sciarretta & Trematerra 2011). To improve on the long-term forecasting of *C. capitata* spatial distributions, in regions like EGVV. A database should be constructed with real-time up to date weather data, including data for at least temperature, relative humidity, rainfall, heat units and cold units. This data could then be used in ML to construct up to date seasonal spatial distribution models for *C. capitata,* which will improve over time, given the capabilities of ML algorithms to learn from the data it is given.

The RF analyses from Chapter 4 needed to be expanded to test the robustness of the RF algorithm to classify *C. capitata* HCSs, by increasing the sample size and variability in the predictor variables. Data from the Warmbokkeveld (WB) were sourced for this purpose (also currently under AW-IPM using SIT), another heterogeneous fruit producing region in the Western Cape, which has different climatic and topographic characteristics from EGVV. *Ceratitis capitata* spatial distribution data from WB and EGVV were combined to test the robustness of the RF algorithm. The geographic drivers of *C. capitata* spatial distribution in WB differed from the drivers in EGVV (Chapter 5). Temperature-related variables, which were highly variable in WB, were also the most prominent drivers of *C. capitata* spatial distributions. This indicates that the drivers of *C. capitata* spatial distributions are to some extent area-specific, but also highlight the importance of the variability of the environment within a region, in driving the spatial distribution of *C. capitata.* In all scenarios tested, the distance of orchards to urban areas, where possible unmanaged fruit fly host fruits were located, played an important role in driving the early season spatial distribution of *C. capitata,*

supporting my findings in Chapter 2 and the suggestions made by numerous authors (see (Myburgh 1956; Israely *et al.* 1997; Barnes 2008; DeVilliers *et al.* 2013; Manrakhan & Addison 2014). Increasing the sample size and the predictor variability did not have a significant negative influence on the model accuracies of the RF models, confirming the robustness of the RF algorithm to model complex ecological interactions (Prasad *et al.* 2006; Cutler *et al.* 2007; Olden *et al.* 2008). This study confirmed the importance of long-term environmental variables in determining the spatial distribution of fruit flies. It also highlighted that *C. capitata* spatial distributions are driven by different factors depending on the region. This study further demonstrated that the results of region-specific analyses and the methodology developed and demonstrated therein, can easily be extended to other (and larger) areas. This is indicative that these RF classification models are universally applicable and that scale of analysis is important when area-wide fruit fly managers interpret *C. capitata* spatial distributions.

In conclusion, this study provides the deciduous fruit industry of the Western Cape with a toolset to quantify the spatio-temporal distributions of *C. capitata,* on an area-wide scale, but also on orchard level. The information gained improves our understanding of how *C. capitata* spatio-temporal distributions are driven by stable geographic variables within heterogeneous fruit producing regions. This information will assist the deciduous fruit industry in improving its management of *C. capitata*, by more precise application of fruit fly management actions. It also paves the way for future research on other economically important fruit flies (*Bactrocera dorsalis*) and moth pests (*Thaumatotibia leucotreta*), which are currently threatening the deciduous fruit industry in South Africa. The application of GIS and ML holds promise for the advancement of the agricultural sector, especially in the field of crop protection. It is recommended, that this research be extended to include more precise and up to date pest monitoring data, and that an effort should be made to establish a real-

time weather database for the Western Cape. Furthermore, real-time data on the cultivar phenology is also important. Although this type of data is easily obtainable on an orchard level, the challenge is to obtain this information on an area-wide scale, for entire production regions. Having access to real-time *C. capitata* trapping data, weather data, cultivar information and cultivar phenology, would enable the establishment of real-time seasonal spatio-temporal distribution models (Sciarretta & Trematerra 2014), not only for *C. capitata,* but many other pests, whose spatio-temporal distributions are mainly driven by these factors. Although there is currently a coordinated approach toward management of *C. capitata* in the Western Cape, data capturing and data handling are still aspects which can enjoy more resources, as most management decision-information is locked up in the data. Up until this study, area-wide *C. capitata* trapping data has not been used to its fullest potential, in order to improve the management of this pest. This study created a platform for further research in this field. The use of GIS and ML should be further explored, especially for the prevention and management of new invasive species like *B. dorsalis.*

References

BARNES, B. 2008. Why have fruit flies become so troublesome? *South African Fruit Journal* **7**(4): 56–57.

CORACINI, M., BENGTSSON, M., LIBLIKAS, L. & WITZGALL, P. 2004. Attraction of codling moth males to apple volatiles. *Entomologia Experimentalis et Applicata* **110**(1): 1–10.

COX, J.S.H. & VREYSEN, M.J.B. 2005. Use of Geographic Information Systems and Spatial Analysis in Area-Wide Integrated Pest Management Programmes that Integrate the Sterile Insect Technique. In: Dyck, V.A., Hendrichs, J. & Robinson, A.S. (Eds.) *Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management*. 453–477. Springer, Dordrecht.

CUTLER, D.R., EDWARDS, T.C., BEARD, K.H., CULTLER, A., HESS, K.T., GIBSON, J. & LAWLER, J.J. 2007. Random Forests for Classification in Ecology. *Ecological Society of America* **88**(11): 2783–2792.

DEVILLIERS, M., MANRAKHAN, A., ADDISON, P. & HATTINGH, V. 2013. The Distribution, Relative Abundance, and Seasonal Phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology* **42**(5): 831–40.

DRAKE, V.A. 1994. The influence of weather and climate on agriculturally important insects: an Australian perspective. *Australian Journal of Agricultural Research* **45**: 487.

DUYCK, P.F., DAVID, P. & QUILICI, S. 2006. Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *The Journal of Animal Ecology* **75**(2): 518–526.

HANIOTAKIS, G.E. 1974. Sexual Attraction in the olive fruit fly, *Dacus oleae* (Gmelin). *Environmental Entomology* **3**(June): 82–86.

ISRAELY, N., YUVAL, B., KITRON, U. & NESTEL, D. 1997. Population fluctuations of adult Mediterranean fruit flies (Diptera: Tephritidae) in a Mediterranean heterogeneous agricultural region. *Environmental Entomology* **26**(6): 1263–1269.

KAPATOS, E. & FLETCHER, B.S. 1983. Seasonal changes in the efficiency of McPhail traps and a model for estimating olive fly densities from trap catches using temperature data. *Entomologia experimentalis et applicata* **33**(1): 20–26. Wiley Online Library.

KATSOYANNOS, B.I., KOULOUSSIS, N.A. & CAREY, J.R. 1998. Seasonal and annual occurrence of Mediterranean fruit flies (Diptera: Tephritidae) on Chios Island, Greece: differences between two neighboring citrus orchards. *Annals of the Entomological Society of America* **91**(1): 43–51.

KITRON, U. 1998. Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. *Journal of Medical Entomology* **35**(4): 435–445.

MANRAKHAN, A. & ADDISON, P. 2014. Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science* **70**(4): 651–660.

MCINNIS, D.O., TAM, S., GRACE, C. & MIYASHITA, D. 1994. Population Suppression and Sterility Rates Induced by Variable Sex Ratio, Sterile Insect Releases of *Ceratitis capitata* (Diptera: Tephritidae) in Hawaii. *Annals of the Entomological Society of America* **87**(2): 231–240.

MIDGARDEN, D.G., LIRA, E. & SILVER, M. 2014. Spatial Analysis of Tephritid Fruit Fly Traps. In: Shelly, T., Epsky, N., Jang, E.B., Reyes-Floris, J. & Vargas, R. (Eds.) *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. 277–320. Springer, Dordrecht.

MYBURGH, A.C. 1956. Bionomics and control of the fruit flies, *Ceratitis capitata* (Wied.) and *Pterandrus rosa* (Ksh.), in the Western Cape. University of Stellenbosch.

NEUENSCHWANDER, P. & MICHELAKIS, S. 1979. McPhail trap captures of *Dacus oleae* (Gmel.)(Diptera, Tephritidae) in comparison to the fly density and population composition as assessed by sondage technique in Crete, Greece. *Mitteilungen der Schweizerischen Entomologischen Gesellschaft* **52**(4): 343–357.

NYAMUKONDIWA, C. & TERBLANCHE, J.S. 2010. Within-generation variation of critical thermal limits in adult Mediterranean and Natal fruit flies *Ceratitis capitata* and *Ceratitis rosa*: Thermal history affects short-term responses to temperature. *Physiological Entomology* **35**(3): 255–264.

OLDEN, J.D., LAWLER, J.J. & POFF, N.L. 2008. Machine Learning Methods Without Tears: A Primer for Ecologists. *The Quarterly Review of Biology* **83**(2): 171–193.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I., CAREY, J.R. & KOULOUSSIS, N.A. 2001. Seasonal and annual occurrence of the Mediterranean fruit fly (Diptera: Tephritidae) in northern Greece. *Annals of the Entomological Society of America* **94**(1): 41–50.

PAPADOPOULOS, N.T., KATSOYANNOS, B.I. & NESTEL, D. 2003. Spatial Autocorrelation Analysis of a *Ceratitis capitata*, (Diptera: Tephritidae) Adult Population in a Mixed Deciduous Fruit Orchard in Northern Greece. *Environmental Entomology* **32**(2): 319–326.

PRASAD, A.M., IVERSON, L.R. & LIAW, A. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **9**(2): 181–199.

ROBACKER, D.C., MORENO, D.S. & WOLFENBARGER, D.A. 1990. Effects of trap color, height, and placement around trees on capture of Mexican fruit flies (Diptera: Tephritidae). *Journal of Economic Entomology* **82**(2): 412–419.

SCIARRETTA, A. & TREMATERRA, P. 2011. Spatio-temporal distribution of *Ceratitis capitata* population in a heterogeneous landscape in Central Italy. *Journal of Applied Entomology* **135**: 241–251.

SCIARRETTA, A. & TREMATERRA, P. 2014. Geostatistical Tools for the Study of Insect Spatial Distribution: Practical Implications in the Integrated Management of Orchard and Vineyard Pests. *Plant Protection Science* **50**(2): 97–110.

VERA, M.T., RODRIGUEZ, R., SEGURA, D.F., CLADERA, J.L. & SUTHERST, R.W. 2002. Potential geographical distribution of the Mediterranean fruit fly, Ceratitis capitata (Diptera: Tephritidae), with emphasis on Argentina and Australia. *Environmental Entomology* **31**(6): 1009–1022.

VREYSEN, M.J.B. & SALEH, K.M. 2001. Long-term sampling of gamma sterilised male *Glossina austeni* (Diptera: Glossinidae) with sticky panels on Unguja Island (Zanzibar). *Acta Tropica* **80**(1): 29–37.